



**The International Congress for  
global Science and Technology**



**ICGST International Journal on Graphics, Vision  
and Image Processing (GVIP)**

**Volume (16), Issue (I)  
June 2016**

**[www.icgst.com](http://www.icgst.com)  
[www.icgst-amc.com](http://www.icgst-amc.com)  
[www.icgst-ees.com](http://www.icgst-ees.com)**

**© ICGST, 2016  
Delaware, USA**

GVIP Journal  
ISSN Print 1687-398X  
ISSN Online 1687-3998  
ISSN CD-ROM 1687-4005  
© ICGST 2016



## Table of Contents

Papers	Pages
P1151550454, Gh.S. El-Tawel and Ashraf.K. Helmy , "3-D Object Extraction from Point Cloud LiDar Scan Data"	1--9
P1151549452, Ayman H. Nasr and Hind Z. Abdelhamid, "Identifying Polarimetric Signatures for Different Features in Radarsat-2 PolSAR Image of Part of Halayib Area, EGYPT",	11--18
P1151531394, Nidhi Goel and Ekta Walia, "Fast Content Based Image Retrieval Using Zernike Moments",	19--25
P1151135788, Hao-Hung Tsai and Chiun-Li Chin and Yung-Chih Cheng, "Automatic Pulmonary Embolism Detection System ",	27--36
P1151550456, Anuradha.S.G and K.Karibasappa and B.Eswar Reddy, "A SEGMENTATION SCHEME BASED ON UNIFORM LBP AND MORPHOLOGICAL APPROACH",	37--45
P1151549451, Samia Snoussi and Fethi Ghazouani and Yosra Wahabi, "Text lines Segmentation of Handwritten Arabic Script using Outer Isothetic Cover",	47--55



**ICGST International Journal on Graphics, Vision and Image Processing  
(GVIP)**

**A publication of the International Congress for global Science and Technology -  
(ICGST)**

**ICGST Editor in Chief: Dr. rer. nat. Ashraf Aboshosha**

**[www.icgst.com](http://www.icgst.com), [www.icgst-amc.com](http://www.icgst-amc.com), [www.icgst-ees.com](http://www.icgst-ees.com)**

**[editor@icgst.com](mailto:editor@icgst.com)**



## 3D Object Extraction from Point Cloud LiDar Scan Data

<sup>1</sup>Gh.S. El-Taweel and <sup>2</sup>Ashraf.K. Helmy

<sup>1</sup>Department of Computer Science,

Faculty of Computers and Informatics, Suez Canal University, Ismailia, Egypt

<sup>2</sup>Data Reception, Analysis and Receiving Station Affairs Division,

National Authority for Remote Sensing and Space Sciences, Cairo, Egypt

ashraf.khaled@narss.sci.eg; ghada@ci.suez.edu.eg

### Abstract

Segmentation followed by extraction of a 3D object for LiDar (Light detection and ranging) data is an essential step in many applications, namely, hazard mapping, rock fall, data classification. This paper presents an approach for 3D object segmentation from dense LiDar point cloud data. The concept of the proposed technique is based on statistical analysis over the derivative domain of a point cloud. Initially the data are divided into background and non-ground regions. Then, the non-ground area is further segmented into distinct 3D objects using statistical information. Experimental results in a hilly area show that the proposed algorithm can efficiently segregate different rock units. Comparing with other techniques, it offers the highest rate of correctness, fullness and rapidness for rock unit separation.

**Keywords:** LiDar; point cloud; segmentation; 3D object extraction.

### Nomenclature

LiDar	Light Detection and ranging
Var	Variance
RMSE	Root Mean Square Error
NEO	Number of Extracted Objects
AEB	The Accuracy of Extracted Boundary
CO	Completeness of an Object

### 1. Introduction

Partitioning the data and image segmentation are considered as nonstop challenges in the fields of image processing, computer vision and human visual perception. Change detection, object tracking, image indexing, image classification, motion estimation and others are examples of applications that consider the segmentation process as a mandatory step in their approaches. Nowadays, there are many sources of digital images ranges from photography to satellite images, high and low resolution images, 2D and 3D data, optical and radar images, multi and hyper-spectral images, and SAR Interferometry and LiDar data. This variation of input source makes the process of

designing a segmentation approach is a tricky task. Basically the segmentation process based on either spectral characteristic of the image (pixel value) or spatial resolution of the image (the degree of clearance of the object within the image) or the image texture information (the relation between neighbored pixels). In the field of the remote sensing and 3D images, building 3D objects has extreme importance in different applications, hazard mapping [1], high resolution image classification [2], rock fall [3], and mobile mapping [4]. Many researchers have been able to segment and reconstruct 3D objects from remotely sensed data [5, 6, 7, 8]. Light Detection and Ranging (LiDar) has become an attractive choice for these applications, as it provides a significant information with respect to the x, y and z coordinate of the different object. Through this research, we introduce a segmentation approach for 3D LiDar point clouds. The rest of this paper is reorganized as follows Section 2 presents a brief introduction of light detection and ranging (LiDar) and the concept of the point cloud. Section 3 focus on the point cloud segmentation algorithms, the most recent approaches and its application in remote sensing community. Section 4 presents a propose method of point cloud data segmentation. The evaluation criteria are listed in section 5. Section 6 summarizes the results. Finally the conclusions achieved is summarized in section 7.

### 2. Light Detection and Ranging (LiDar)

The LiDar is an acronym of light detection and ranging, in the remote sensing community, it is used in the applications of generating high resolution maps, archaeology, forestry, geology, seismology. LiDar system has its own source of energy, it transmits a low-power laser pulse to illuminate a target and then the reflected back energy is recorded by the system. The LiDar system can precisely identify the ground location (x, y, z) of the target object with an accuracy of sub-centimeters. This is achieved by measuring the time spent of the emitted pulse to travel to the object and back to the system [9]. Laser pulse consists of photons that travel in space with light speed, it travels towards the object, part of the



incident light hit the object and returns back to the sensor. If the size of the object is small (such as a branch of a tree), some of the incident light will pass through and continue travelling to the ground as shown in figure 1. This leads to a multiple reflections from one laser pulse. According to this transmitting/receiving mechanism the backscatter of a transmitted pulse will be recorded as a waveform (distribution of backscattered pulse) as shown in figure 1. Nowadays, there are three different LiDAR scanning technologies [9]:

- i. A technique based on measure the time interval between transmitting the pulse and record its backscatter, from which the distance to the object can be calculated.
- ii. A technology based on emitting a series of pulses simultaneously with a known phase, then the system calculates the distance to the object by comparing the returned phase to the incident phase.
- iii. The last technique use two mounted sensors to hit the same object simultaneously, then comparing the returned pulse to calculate the required distance.

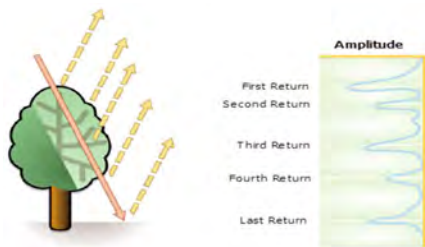


Figure 1. An example of LiDAR waveform. © ESRI- CO. ARC GIS

### 3. Points cloud segmentation

Image segmentation in LiDAR data aims to identify and reconstruct the 3D object, and separating it from the background and other features. The most tradition techniques in image segmentation are image thresholding, the threshold value determines the boundary between an object and its background, commonly the threshold value is determined by histogram analysis.

Mass, 1999 used the height's texture information for automatic segmentation and detection of objects like tree, roads and urban area [10]. Brunn and Weidner, 1997, used heights information derived from Lidar imagery for building extraction; all objects above the ground are detected using direct threshold value, then buildings are detected by applying parametric building model [11]. Felzenszwalb and Huttenlocher introduced an image segmentation algorithm for 2D image based on a graph representation [12]. Firstly, they represent the image as a graph, two different types of graphs are used in their study. A local neighborhood of image pixels is used to build the graph nodes, then they measure the color difference between neighbor's nodes. The other method uses the spatial coordinate (x, y) side by side with pixel color to construct the graph. Then they seek the evidence for the existence a boundary between two regions. The evidence is based on both of intensity difference across the claimed

boundary and the intensity difference within each segment. Triebel et al. extend the Felzenszwalb segmentation algorithm to suit an indoor laser scanning [13], he used a graph-based clustering techniques to segment the range images. He built a graph for both feature and geometric space. The first graph is used to preliminary divide the range image into different regions, while the other one is used to smooth the segment classes in 3D.

Orthuber and Avbelj introduced a workflow approach to detect and reconstruct the building from a Lidar point cloud based on region growing method [14]. He succeeds to extract a building roof, the step and intersection edges can be delineated between roof segments.

Victor et. al. introduced a method to segment an individual tree crowns from airborne Lidar data [5]. They proposed a method based on using the topological structure of the forest. The point cloud arranged in hierarchical data structures, the relationships of tree crown components are used as a weight factor in the graph representation. Finally the individual tree crowns are separated.

Our proposed method of point cloud segmentation based on shape feature calculated using the geometric gradient of the neighboring points and a normal vector at a surface for each point cloud. In the following subsections, the concept of point gradient and surface normal are presented.

#### 3.1 Point cloud Gradient

In the field of signal processing, an image gradient is defined as the rate of change with respect to the color or intensity. Usually image gradient is represented as a vector with amplitude and phase. The variation through the image can be represented from high to low value or from white to black as shown in figure 2. The image gradient is used in many image processing and pattern recognition applications such as, edge detection, color blending, and image fusion.

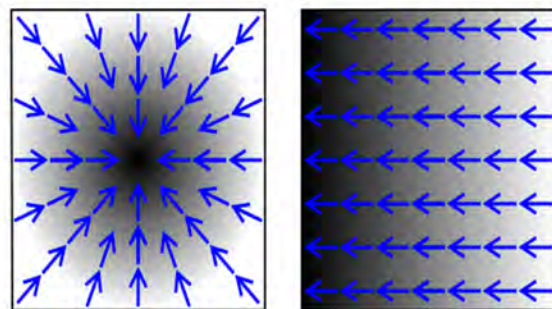


Figure 2. Two types of gradients, with blue arrows to indicate the direction of the gradient. © Wikipedia

The mathematical form of an image gradient of an image is a 2 D vector. The vector is defined at each point of the image as the derivatives in horizontal and vertical directions. It represents the direction of intensity growing. In our case, we extend the use of the image gradient of a point cloud in a 3D domain, where the intensity variable is replaced with a point spatial location (x, y, z). As the gradient vector represents the rate of change in point slop, the gradient of a point cloud is calculated as:

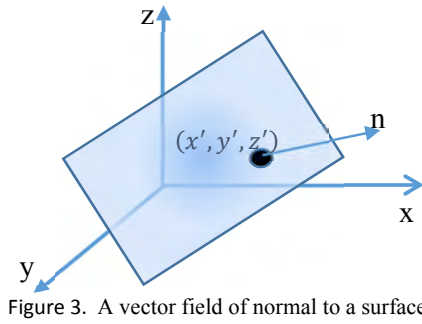


$$\nabla f(x, y, z) = \text{grad}(f) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix} \quad (1)$$

Where,  $\frac{\partial f}{\partial x}$ ,  $\frac{\partial f}{\partial y}$  and  $\frac{\partial f}{\partial z}$  are the gradient at x, y and z direction respectively?

### 3.2 Surface normal

Normal of an object is defined as a unit vector perpendicular to that object at a given point. For 2D signals, the normal to spline at a certain point is a vector with unit length that is orthogonal to the tangent line at that point. This definition is extended to the 3D as follows: “the normal to a surface at given point is defined as a unit vector perpendicular to the tangent plane at that point”.



As shown in figure 3, for a given surface S,

$$S = f(x, y, z) = ax + by + cz + d = 0 \quad (2)$$

The normal vector is defined as:

$$n = \nabla f = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (3)$$

Where  $\nabla f$  is a gradient

The normal vector passing through a point  $(x', y', z')$  is defined as:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} x - x' \\ y - y' \\ z - z' \end{bmatrix} = a(x - x') + b(y - y') + c(z - z') = 0 \quad (4)$$

## 4. Proposed segmentation method of Lidar point cloud.

The system flowchart of a proposed approach is shown in figure 4. The Lidar cloud point is firstly segmented into background and non-background regions, then the boundary of each non-background object is delineated. The proposed method for the segmenting the Lidar cloud point based on “voxel grid based segmentation”, and contain the following three main steps:

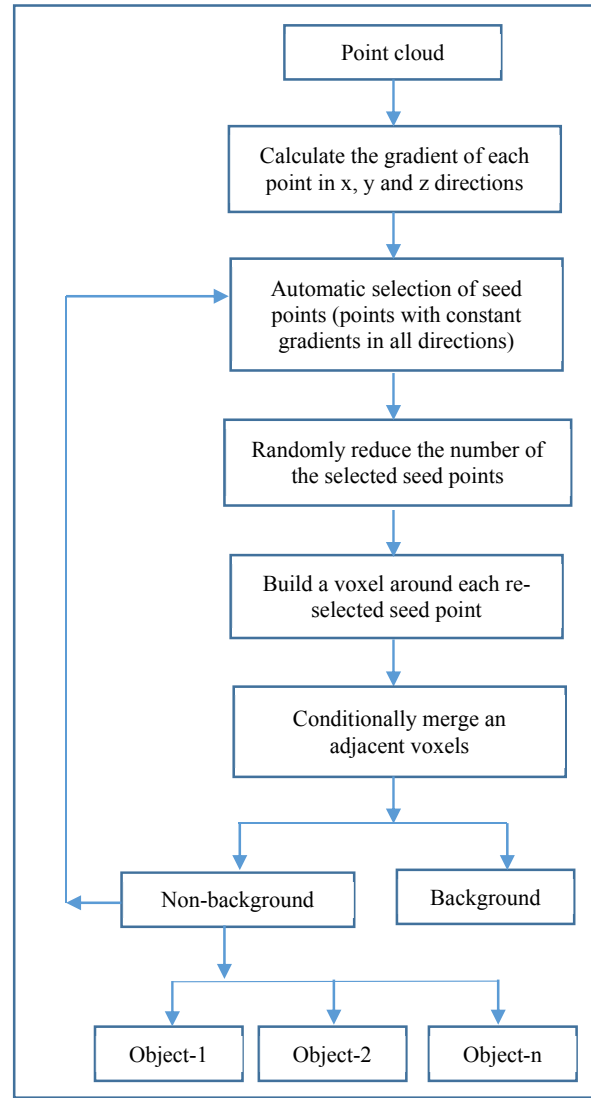


Figure 4. Process model of 3D object reconstruction from LiDAR point cloud

### I. Segment the entire 3D space into many voxels.

Voxel is defined as a unit cube around candidate point (P) in the 3D-space. All points belong to a certain voxel are closer to p than any other point, Japans voxel segmentation method [15] based on firstly, number of the seed points are nominated and considered as a base for a voxels. Then he assumed that all 26-adjacent voxels are contained within constant distance equal to  $\sqrt{3} * R_{\text{voxel}}$ . Where  $R_{\text{voxel}}$  and  $R_{\text{voxel}}$  are the seed and voxel resolution respectively, it is a user defined values and depend mainly on the nature of the objects. The main problem of this method is the selection criteria of both voxel and seed resolution. Our proposed method for selecting a seed point depends on calculating the gradient for each point as described in section 3.1. Firstly, we define the neighbored points (PN) around a candidate point (P) in the 3D-space as the all points satisfy the following conditions:

$$P_N \Leftrightarrow |f_x| = C_1 \ \&\& \ |f_y| = C_2 \ \&\& \ |f_z| = C_3 \quad (5)$$





Where  $f_x, f_y$  and  $f_z$  are the gradient at x, y and z directions respectively.  $C_1, C_1$  and  $C_1$  are small constant values approaches zero, this ensures that the selected point lies on the object surface not in its edge.

Then, we build a smooth cube surface around a seed point, according to the following iterative procedures:

In each iteration, the point that will be included in the cube is the point that has constant variations with respect to the surrounded points. The constant variation is defined in terms of the gradient in all directions (equation 1) and the unit vector perpendicular to the surface at the point P (equation 4). The normal to a surface at a point P is defined as a vector perpendicular to the tangent plane to that surface. Table 1 presents a Pseudo code of the proposed method for selecting a seed point and building a cube around it.

## II. Background extraction.

The background is delineated by merging adjacent voxels based on some statistical measurements. Neighbored voxels are grouped together if they meet predefined requirements. Three categories of adjacency are introduced in the literature, 26, 18 or 6. Two voxels are considered 26 adjacent if they are shared in a vertex or an edge or a face. On the other hand it said that the two voxels are 18 adjacent if they share an edge or a face, while 18 adjacent comes to the scene if they share only a face. In our experiment we consider the 26 adjacent explanation. We will consider the grouped background voxels as a separator between different objects in space. This idea is first introduced by Douillard et. Al, he used the local neighbored as the only parameter to perform the separation process [16], we extend this method as follows: Consider the candidate voxel  $v$  and the 26-adjacent voxels are  $N_i$ , it is said that  $N_j$  belongs to  $v$  according to the following condition:

$$var_x = var_y \mid |var_x = var_z| \mid var_y = var_z \quad (6)$$

$var$  is defined as the variance of  $N_j$  voxel. The variance of a voxel  $i$  is calculated as follows:

$$var = \sum_{i=1}^n p_i \cdot (x_i - \mu) \quad (7)$$

Where:

$x_i$  represents a point cloud in a voxel  $v$ .

$n$  is the number of points belong to a voxel.

$p_i$  denotes the probability mass function of voxels point.

This is a guarantee that the continuity of an object and ensure that the boundaries do not break up.

## I. Object extraction.

The boundary of each non-background voxels (object-based) is addressed using the difference between neighbored normal for each voxel (equation-4). Then each point cloud fit to the voxels edge (satisfy the condition pronounced in equation 6) is differently marked.

## 5. Evaluation criteria for segmentation process.

The performance evaluation of the segmentation results is considered by direct comparison with different segments

determined by visual interpretation of the LiDAR point cloud. We follow the evaluation procedures proposed by [17, 18].

Table 1 Pseudo code of proposed method

---

```

Input: point cloud; point cloud
Output: smooth cube around a selected seed point
Begin:
// select a seed point
For all point in the 3D Space do
{
     $|f_x|, |f_y|, |f_z| \leftarrow$  Calculate the gradient in x,
                                y, z directions of each
                                point in the entire
                                space as described in
                                Equation -1 and a
                                selection criteria
                                introduced in equation-
                                5
    If  $|f_x| = Constant \approx 0$ ;
    | If  $|f_y| = Constant \approx 0$ ;
    | | If  $|f_z| = Constant \approx 0$ ;
    | |  $P_s \leftarrow$  Select the point as
a seed point;
    | End;
    | End;
    End;
}
// Build a smooth cube around a seed point.

For all point  $\in P_s$  do
{
    For all point N around a selected seed
    point;
    {
        Calculate the normal  $n_i$  at each
        point as described in equation 4;
         $n_i = \nabla f = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ ;
        for i  $\leftarrow 0$  to N do
        {
            If  $\Delta n_i = 0$ ;
            Assign a point to a cube
            around a selected point ;
        }
    }
}
End;
```

---

In our evaluation process, we address three types of evaluations; with respect to the number of extracted object, the accuracy of extracting boundary, and completeness of an object. The general concept of the used criteria is illustrated in figure 5. The segment marked "A" is the reference segment, It is identified by examining the point cloud from various perspective angles. The segment "B" represents imperfect segments comes out from segmentation algorithm. In this procedure, we define two types of errors as follows:





*errr-1* is defined as the count of all numbers of points cloud in the reference segment that not exist in the imperfect segment (it is shown in red color), while *errr-2* identify the imperfect segment that has not an identical pair in the reference segment (appear as blue color). In general, the segmentation quality is represented by the total error, Total number of *error* = *errr-1* + *errr-2*.

This error represents “the correctness and completeness of an object”. Comparing the number of the objects in segmented scene, gives an estimate of the completeness of the segmentation process. On the other hand, the accuracy of the extracted boundaries is addressed by calculating RMSE with respect to the manually detected reference object, and calculated as follows:

$$RMSE = \frac{1}{N} \sum_{i=1}^N (p_{ref(i)} - p_i)^2 \quad (8)$$

Where  $p_{ref(i)}$  is the reference point (i) and N is the total number of points.

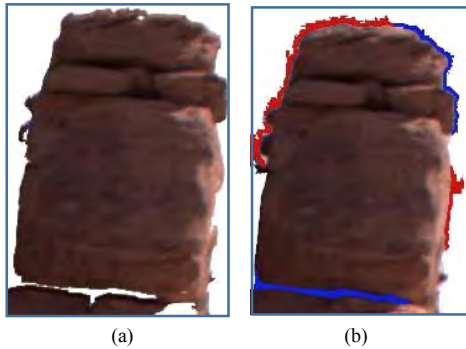


Figure 5 Example of different Type of errors included in performance evaluation: (a) Reference segment (b) faulty segment

Moreover, through our evaluation procedures, we consider a direct comparison with different approaches of 3D segmentation. The performance comparison with other techniques is a difficult task due to the different nature of objects and scanning conditions [19], we compared our algorithm with Jeremie Papo [15] and E. Orthubera [14] due to similar conditions specifically in terms of generality with respect to the target segmented objects, we will refer to these algorithms in succeeding analysis as *alg1* and *alg2* respectively.

Jeremie Papon proposed an approach for point cloud segmentation based on voxel cloud connectivity [15], they make use of both 3D geometry and spectral characteristics obtained from an RGB camera to delineate the boundaries of an object. This algorithm based mainly on seeding algorithm in 3D space and local iterative clustering for classifying RGB image. Orthubera proposes a system work flow for segmentation and extraction a roof building from a LiDar point cloud. This method based on region growing with adaptive threshold [14].

## 6. Experimental results

Different objects from the Lidar point cloud, this step is considered essential in many applications such as, hazard mapping and rock fall detection. To meet this requirement, we use a data set to cover the hilly area in about all-rich

area. Optech ILRIS terrestrial laser scanner was used to collect a point cloud at pulse rate 12.5 kHz with accuracy 0.26 cm, the technical specification of laser scanner is presented in the appendix-A. The experimental results conducted in this research have been executed using Matlab code on processor i7-2.9 Ghz-6 GB Ram.

The raw data consists of 1,425,866 points, the record of each point contains its spatial location (x, y, z) and RGB value. The photographic image of the study area is shown in figure 6. Figure 7 shows the isometric view of the 3D point cloud of the study area, different rock units are expected to be identified and isolated separately.



Figure 6. Photographic image of the study area

Figure 8 demonstrate the output segment scene of a proposed algorithm, in which each segmented unit is assigned to a unique color. The segmentation results of *alg1* and *alg2* are shown in figures 9, 10. Respectively. As mentioned in section 5, we compare our approach with *alg1* and *alg2* to analyze different segmentation results. The best segmentation results are declared according to the handmade segmented data set. The accuracy percentage and computation time are recorded and listed in table-2.

From the visual analysis of the results we can conclude that all algorithms in the middle of the image can be segmented and extracted with various degrees of success in terms of the accuracy of Extracted Boundary and completeness of an object. For instance, consider the green circle in figures 8, 9. The *alg1* introduced two types of objects inside the circle while a proposed method shows strong consistency in delineation an object boundary. This indicates that the *alg1* fails in completeness test. In the same direction we can find that an *alg2* also comes with incomplete completeness with respect to object reconstruction, this is can be shown as red circle in figure 10. On the other hand the *alg1* and *alg2* missed some small rock unit in the left part of the area (pointed with red arrows in figures 9 and 10 while the proposed algorithm correctly identified these units as shown in figure 8. The accuracy of the extracted boundary is a difficult task, so we present a closer look at the results shown in figure 11; it represents a zoomed area of the previous figures, the edges of the extracted objects are highlighted with white color. Comparing the boundaries of the extracted objects from figure 10, it is clear that *alg1* and *alg2* suffers from the non - completeness of object content and non-correctness of object boundary. On the other hand all algorithms, including a proposed method come with a zigzag attitude in the object's edges.



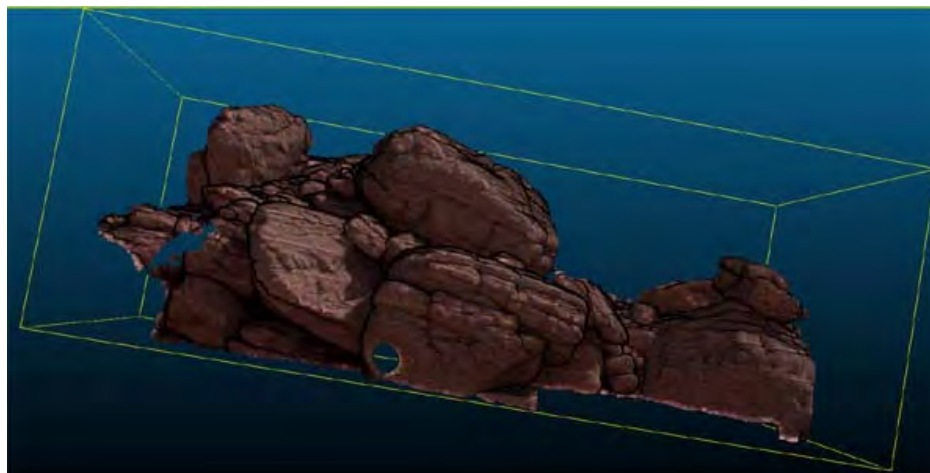


Figure 7. Isometric view of the study area

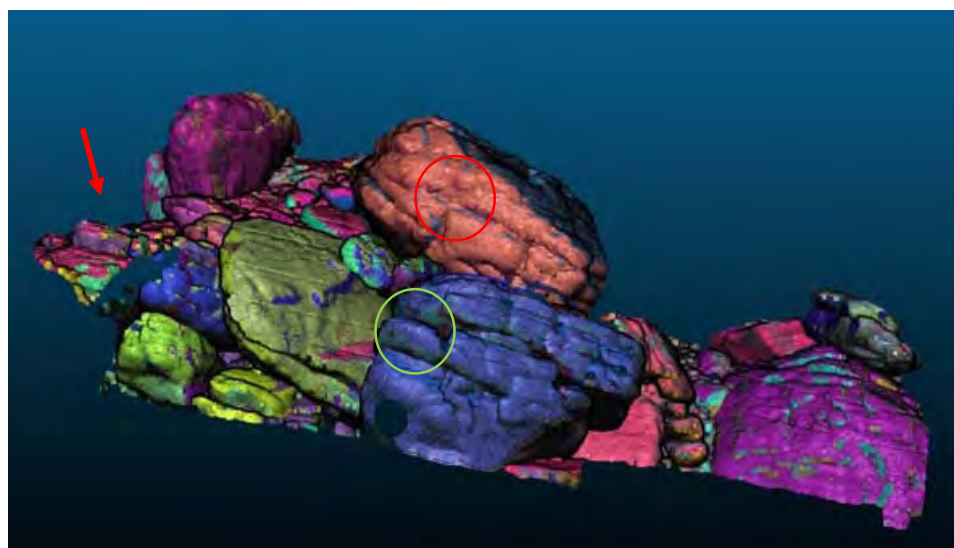


Figure 8. Segmentation results of a proposed method

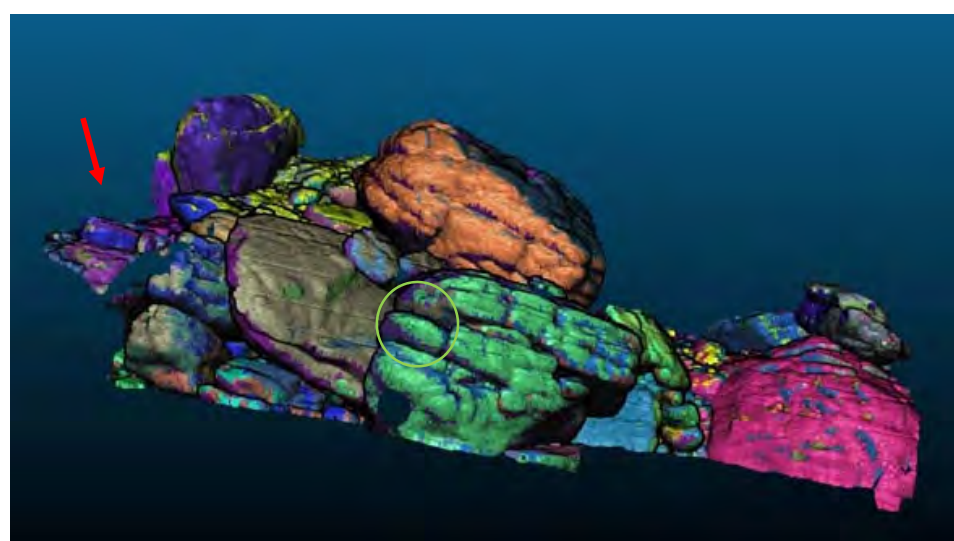


Figure 9. Segmentation results of the alg1



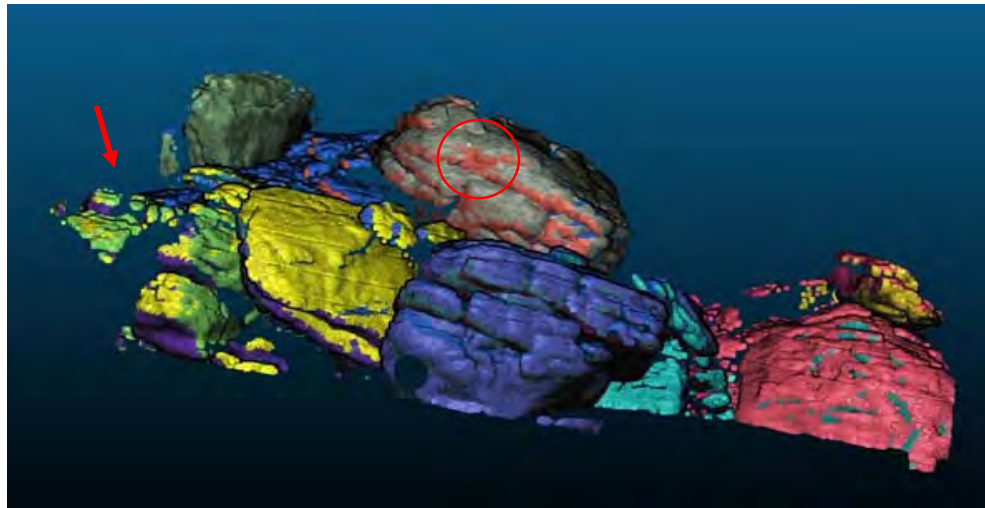


Figure 10. Segmentation results of the alg 2

Quantitative analysis of a proposed method compared with other techniques is tabulated in table 2. The results show that *alg2* comes with the worst results in terms of the number of identified rocks (about 65%) as many of the rocks are misclassified and identified as background, the proposed method achieved the highest accuracy. A proposed method and *alg1* achieve nearly the same accuracy with respect to object completeness with slight improvement of a proposed method (about 5% higher). *alg2* surplus the others in terms of the accuracy of the extracted edge (about 0.3% higher than a proposed method) this because of the *alg2* uses a boundary smoother to fine tune the extracted object.

Table 2. Performance evaluation of a proposed method compared to others; NEO: Number of Extracted Objects, AEB: the Accuracy of Extracted Boundary (RMSE, in meters), CO: Completeness of an Object.

	NEO %	CO%	RMSE	Computation time (Sec.)
A proposed method	98	92	2.02	18
<i>Alg1</i>	82	87	2.39	23
<i>Alg2</i>	65	62	4.15	20

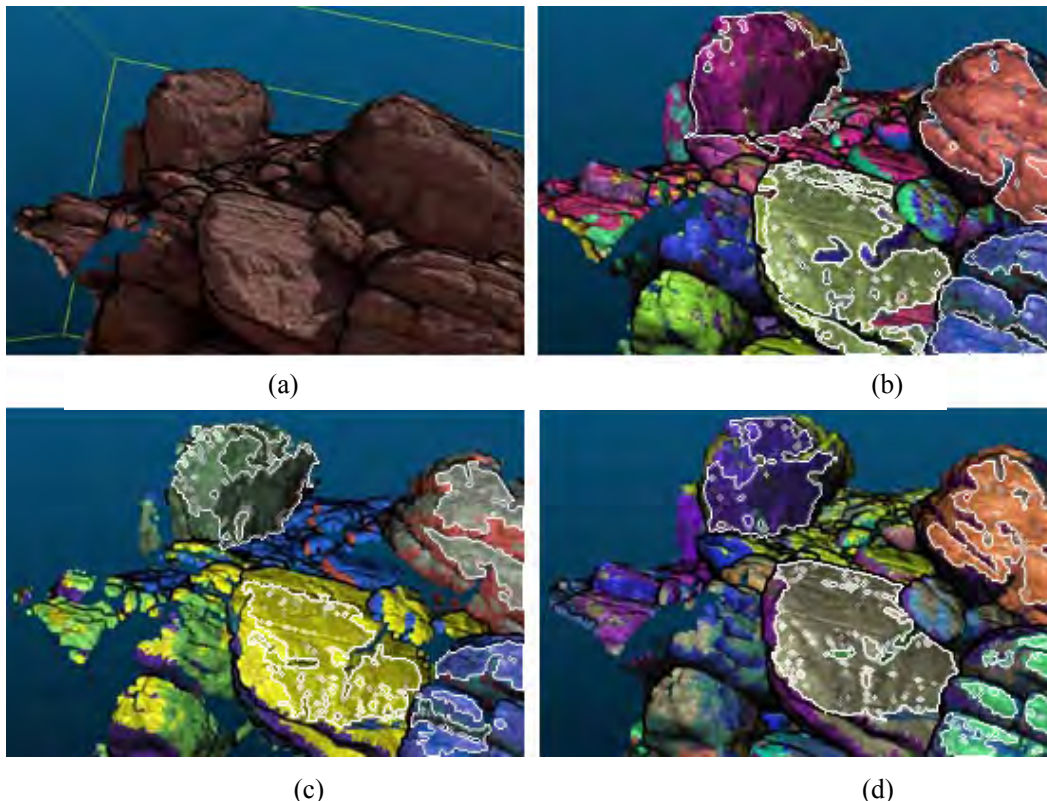


Figure 11. Zoomed area with edge highlighted of selected objects (a) isometric view of the study area; (b) proposed method; (c) output of *alg1*; (d) output of *alg2*





## 7. Conclusions

In this paper, we have proposed a framework of dense LiDAR point cloud data segmentation associated with 3D object extraction. The proposed method makes use of a statistical properties of the individual point and its neighbors in a 3D space such as gradient, variance and the orientation. Firstly, the entire 3D space is portioned into the background and others by using the difference between a gradient of each point and its neighbors. Then the non-ground region is further divided into different objects via calculating the orientation changes of a normal vector to a point plan tangent, it is calculated from a point to the next through the entire space. The experimental results show that the proposed method achieved an efficient rock unit segmentation process and object extraction achievement. Also, it has been shown that the proposed method can separate small scattered rock units. Comparing with other algorithms, the proposed method surplus others in terms of completeness, correctness, segmentation's accuracy and computation time. It is highly recommended to incorporate a boundary smoother of the extracted object, this issue will be addressed in future work.

## 8. Acknowledgements

The authors would like to express their deepest thanks and appreciation to the National Authority of Remote Sensing and Space Sciences (NARSS) for providing the scanned data for this study.

## 9. References

- [1] Brian D. Collins, Ph.D., P.E., M.ASCE and Greg M. Stock, Ph.D., P.G. "Lidar-based rock-fall hazard characterization of cliffs" GeoCongress 2012 © ASCE 2012.
- [2] Sohel Syed, Paul Dare, Simon Jones, "Automatic Classification of Land Cover Features With High Resolution Imagery And Lidar Data", an object-oriented approach" Proceedings of SSC2005 Spatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Sciences Institute, September, Melbourne: Spatial Sciences Institute. ISBN 0-9581366-2-9, 2005.
- [3] T. Oppikofer, M. Jaboyedoff, L. Blikra, M.-H. Derron, and R. Metzger, "Characterization and monitoring of the Arknes rockslide using terrestrial laser scanning" Nat. Hazards Earth Syst. Sci., 9, pp. 1003–1019, 2009
- [4] Tao, C. V. "Advances in mobile mapping technology", Volume 4 of International Society for Photogrammetry and Remote Sensing book series. (ISSN 1572-) Taylor & Francis. ISBN 0-415-42723-1, 2007.
- [5] Victor F. Strîmbu, Bogdan M. Strîmbu, "A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data", ISPRS Journal of Photogrammetry and Remote Sensing vol. 104, pp 30–43, 2015
- [6] Bottai, L., Arcidiaco, L., Chiesi, M., Maselli, F., "Application of a single-tree identification algorithm to LiDAR data for the simulation of stem volume current annual increment", J. Appl. Remote Sens. Vol. 7, 073699, 2013.
- [7] Karen F. West, Brian N. Webb, James R. Lersch, Steven Pothier, Joseph M. Triscari, and A. Evan Iverson, "Context-driven automated target detection in 3D data," in Proceeding of SPIE, Firooz A. Sadjadi, Ed., Bellingham, WA, SPIE, vol. 5426, pp. 133–143, 2004.
- [8] Palenichka, R., Doyon, F., Lakhssassi, A., Zaremba, M.B., "Multi-scale segmentation of forest areas and tree detection in LiDAR images by the attentive vision method", Selected Topics Appl. Earth Observ. Remote Sens., IEEE J. 6, pp. 1313–1323, 2013.
- [9] S. E. Chen. Laser scanning technology. In J. A. M. Rodriguez, "Laser Scanning Technology for Bridge Monitoring", InTech, ISBN 978-953- 51-0280-9, 2012.
- [10] Mass, G.-H., "The potential of height texture measures for the segmentation of airborne laser scanner data", presented at the 4<sup>th</sup> Airborne Remote Sensing Conference and Exhibition, Ottawa, Ontario, Canada, pp. 21-24, June 1999.
- [11] Brunn, A., Weidner, U. "Extracting buildings from digital surface models", IAPRS Vol. 32, part 3-4w2, Stuttgart, 1997.
- [12] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient graph-based image segmentation", International Journal of Computer Vision, vol. 59(2), pp. 167– 181, 2004.
- [13] R. Triebel, J. Shin, and R. Siegwart, "Segmentation and unsupervised part-based discovery of repetitive objects", In Proceedings of Robotics: Science and Systems, Zaragoza, Spain, June 2010.
- [14] E. Orthubera, b, J. Avbelja, "3D Building Reconstruction From Lidar Point Clouds By Adaptive Dual Contouring", ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. II-3/W4, PIA15+HRIGI15 – Joint ISPRS conference, pp. 25–27 March 2015, Munich, Germany, 2015.
- [15] Papon, Jeremie, et al. "Voxel cloud connectivity segmentation-supervoxels for point clouds." Computer Vision and Pattern Recognition (CVPR), IEEE Conference, 2013.
- [16] B. Douillard, J. Underwood, N. Melkumyan, S. Singh, S. Vasudevan, C. Brunner, and A. Quadros "Hybrid elevation maps: 3d surface models for segmentation", In Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2010.
- [17] Rutzinger, M.; Rottensteiner, F.; Pfeifer, N. A comparison of evaluation techniques for building extraction from airborne laser scanning. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., vol. 2, pp. 11–20, 2009
- [18] Rottensteiner, F.; Sohn, G.; Jung, J.; Gerke, M.; Baillard, C.; Benitez, S.; Breitkopf, U., "The ISPRS benchmark on urban object classification and 3D building reconstruction", ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., vol. I-3, pp. 293–298, 2012.



- [19] Vauhkonen, J., Ene, L., Gupta, S., Heinzl, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., Maltamo, M., "Comparative testing of single-tree detection algorithms under different types of forest. Forestry", vol. 85, pp. 27–40, 2012.

## 10. Appendix



Table A-1 Technical Specifications of ILRIS device

Parameter	Optec ILRIS 3D
wavelength	1550 nm
Minimum range	3 meters
Maximum range	1500 m at 80% reflectivity
Average data acquisition rate	2500 points per second
Laser Beam diameter	29 nm @ 100 m
Distance accuracy	7nm @100 m
Position accuracy	8 nm @100m
Angular accuracy	0.00115
Scanner weight	13 kg without batteries

© TELEDYNE Optic ILRIS

## Biographies



information fusion.

Gh. S. El-Tawel received B.S., M.S., in computer Sciences and Ph.D. degree in Information technology from Cairo University in 1996, 2000, and 2005 respectively. She is currently an associate Professor in Suez Canal University, Ismailia, Egypt. Her research interests include image processing, DNA, and



of interest includes signal and image processing.

Dr. Ashraf Khaled Helmy is Associate Research Professor in the National Authority for Remote Sensing and Space Sciences. He is currently the Head of the Space Sciences and strategic studies Division. He has got his Ph.D. in information technology from Cairo University, 2005. His area







## Identifying Polarimetric Signatures for Different Features in Radarsat-2 PolSAR Image of Part of Halayib Area, EGYPT

Ayman H. Nasr and Hind Z. Abdelhamid

National Authority for Remote Sensing and Space Sciences (NARSS), 23 Joseph Tito st.,  
El-Nozha El-Gedida, P.O. Box: 1564 Alf-Mascan, Cairo, Egypt  
[aymanasr@hotmail.com](mailto:aymanasr@hotmail.com), [sar.hend87@gmail.com](mailto:sar.hend87@gmail.com)

### Abstract

In fully polarized SAR (PolSAR) data the returned signal from a target contains all polarizations. More information about this target may be inferred with respect to single-polarization. Distinct polarization separates targets due to its different backscattering responses. A Radarsat-2 PolSAR image acquired on December 2013 of part of Halayib area (Egypt) was used in this study. Polarimetric signatures for various features (Wadi deposits, Tonalite, Chlorite schist, and Radar penetrated areas) were derived and identified. Their Co-polarized and Cross-polarized signatures were generated, based on the calculation of the backscattered power at various ellipticity and orientation angles. Graphical 3D-representation of these features was provided and more details of their physical information are depicted according to their different polarization bases. The results illustrate that polarimetric signatures, obtained due to factors like surface roughness, dielectric constant and feature orientation, can be an effective representation for analyzing various features. The shape of the signature is significant and can also indicate the scattering mechanisms dominating the features response.

**Keywords:** *Radarsat-2 PolSAR, Polarimetric signatures, Pauli decomposition, Co-polarization and cross-polarization signature.*

### Nomenclature

DEM	Digital Elevation Model
EM	Electromagnetic wave
H	Horizontal
V	Vertical
Lin	Linear
PolSAR	Polarimetric SAR
SLC	Single Look Complex
SAR	Synthetic Aperture Radar

### 1. Introduction

Polarimetric SARs provide significantly more data relative to conventional radars that record backscatter only at the linear polarizations. They allow measurement of the physical characteristics by using scattering mechanism between electromagnetic (EM) wave and the targets [1], [2]. They have been also successfully employed to classify and separate a wide range of terrain types. Fully polarimetric radars record the complete four coherent channels (HH, VV, HV, and VH) and retain the phase information. The “H” indicates horizontal and “V” indicates vertical transmit or receive polarization. The phase differences can result from a time delay when the phase velocity of H and V waves differs within the target. The total power is the sum of the power recorded for each of the linear polarizations (HH, VV, HV, and VH) [3].

Polarimetric signature plot is a general approach to visualize the signature that captures many scattering characteristics of the ground cover targets. It is a 3D-representation of polarimetric information in various polarization bases [4]. It is usually displayed assuming: the identical transmit and receive polarizations (co-polarized) and the orthogonal transmit and receive polarizations (cross-polarized) of the wave intensity at all possible ellipticity and orientation angles. The shape of these plots is significant and can indicate the scattering mechanisms (surface, double-bounce, or multiple/volume) dominating the target response. Ellipse geometric elements are two dimensions, and target response is the third dimension represented in three dimensional coordinate system [5]. The shape of the Polarization signatures of the same target observed in different time should resemble each other in general if there are no changes; otherwise, they should be different. In





some researches [6], [7], [8], and [4] it has been used as a tool for analysis and assessment of various targets.

Pedestal height is an indicator of the presence of an unpolarized scattering component, and thus the degree of polarization of a scattered wave. It can be derived and visualized on the three dimensional polarization signature plots generated from fully polarimetric data [8]. The minimum intensity indicates the pedestal height of the polarization signature. The co-polarization pedestal height is the ratio of the maximum to the minimum received intensity when the polarizations of the transmitting and receiving antenna are the same [3]. Signatures with significant pedestals are typical of targets that are dominated by volume scattering or multiple surface scattering. Van Zyl [9] and Ray et al. [10] found that pedestal height was related to surface roughness with increases in roughness resulting in higher pedestals.

The remainder of the paper is structured as follows: Section 2 presents the acquired data and the study area. Section 3 focuses on the full description of a used method, including the main concepts of scattering matrix, coherency matrix, speckle filter,

geometric correction, feature extraction, Pauli decomposition, and the polarimetric signature. The main results and discussions are outlined in section 4. Finally, Section 5 summarizes the conclusions achieved by this study.

## 2. Acquired Data and the Study Area

A Radarsat-2 PolSAR image acquired on December 2013 was used in conducting this study. It was delivered as a Single-Look Complex (SLC) Standard Quad Polarization, Q6 in compressed format. The major characteristics of the image are depicted in Table 1. The study site selected is part of Halayib area located in south eastern desert of Egypt, with coordinate of  $22^{\circ} 29'$  to  $22^{\circ} 09' N$ , and  $35^{\circ} 43'$  to  $36^{\circ} 03' E$ , as shown in figure 2. A variety of different features were considered in this study. These include (Wadi deposits, Chlorite schist, Tonalite, and Radar penetrated areas). Reference data was collected from geological map and ETM-8 image to verify the identification of these features. The Software used was the freeware Polarimetric SAR Data Processing and Educational Toolbox; (PolSARpro). Figure 1 depicts the snapshot of the PolSARpro software to get the polarimetric signature results.

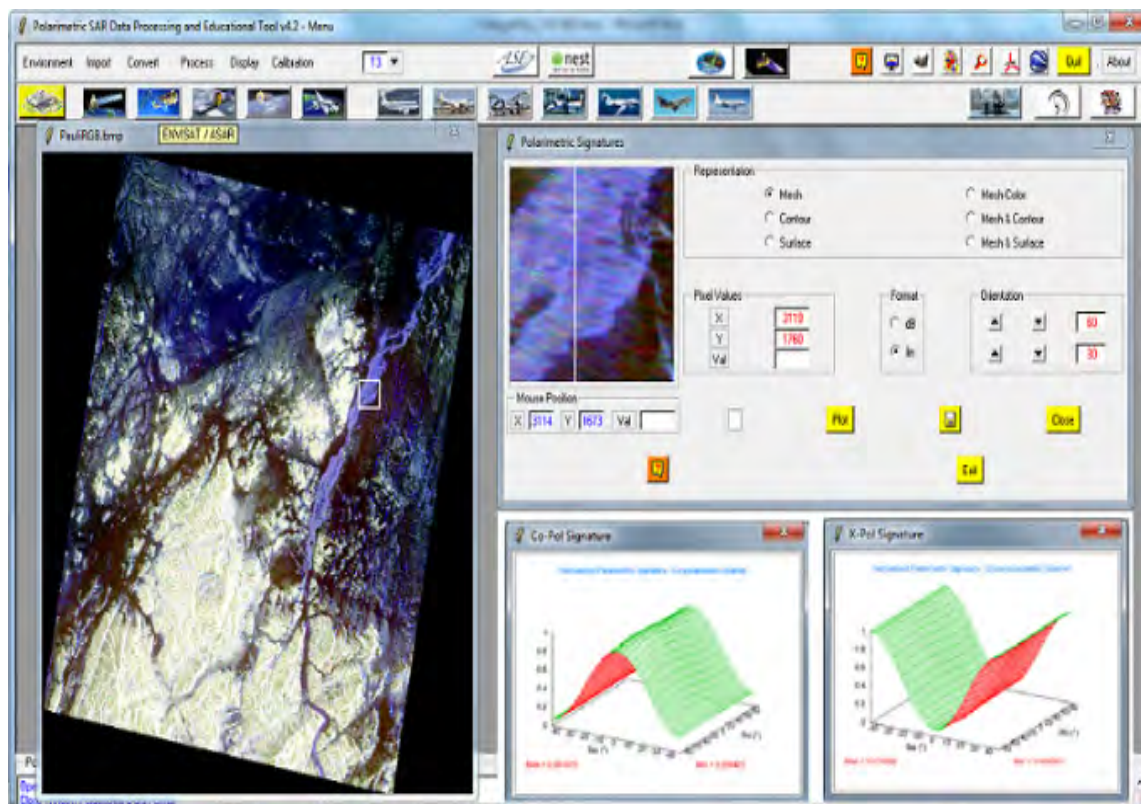


Figure 1: Snapshot of the PolSARpro software to get the polarimetric signature results.



Table 1: Major Characteristics of the Used Radarsat-2 PolSAR Image

RADARSAT-2 (SLC) Image	
Scene Date and time	18-12-2013, 03:38:58
Acquisition Type, beams	Standard Quad Polarization, Q6
Polarizations and Pass-Direction	HH VV HV VH & Descending
Incidence Angle (units: deg)	24.5
Bits Per Sample ( Real, Imaginary)	16, 16
Number of Column and Lines	1465, 7039
Sampled Pixel Spacing (units: m)	7.987
Sampled Line Spacing (units: m)	4.700



Figure 2: The study area (yellow rectangle)

### 3. Methodology

In order to interpret and retrieve the feature information of polarimetric SAR data, pre-processing is of critical importance [11]. The first step is the generation of  $2 \times 2$  scattering matrix  $[S]$  that measures the complete information of the surface features. This is followed by deriving the  $3 \times 3$  coherency matrix  $[T3]$  and the polarimetric parameters. Once the scattering matrix and the covariance matrix are known, one can synthesize the received power for any transmit and receive antenna polarizations. Finally, speckle filtering and geometric correction (geo-referencing) are conducted for interpreting the image correctly.

After completion of this phase, the coherency matrix is decomposed based on the Pauli basis for deriving the polarimetric signature for each feature on the study area. Figure 3 shows the flowchart of the different pre-processing steps for polarimetric signature retrieval.

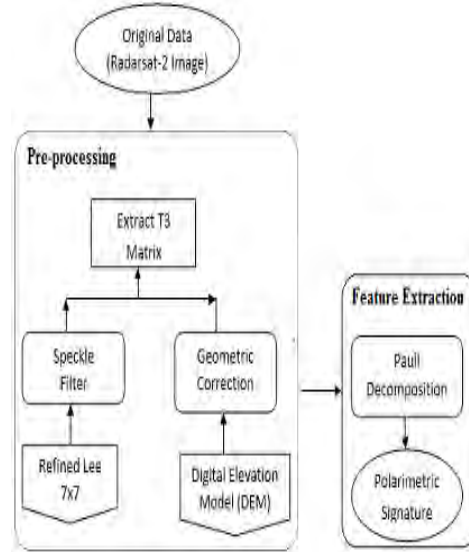


Figure 3: Flowchart of the different pre-processing steps for polarimetric signature retrieval.

#### 3.1 The Scattering Matrix

The original data is in Single Look Complex (SLC) format. When the incident radar signal interacts with the earth feature on horizontal or vertical wave, the backscatter of the radar signal is the contribution of both vertical and horizontal wave. Therefore, the reflected backscatter can be represented by scattering matrix as given in equation 1:

$$[S] = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad (1)$$

Each pixel in a polarimetric radar image is represented by this 4- component scattering matrix. Each component is a complex value which has magnitude and phase of 4- polarimetric channels. The diagonal elements of the scattering matrix are the co-polarized information while off- diagonal elements represent the cross- polarized information [12].

#### Coherency Matrix

Scattering matrix is used to represent the backscatter of the coherent or pure target like urban area. In contrast, the natural target which partially polarized waves is very difficult to be realized using scattering matrix. To describe the distributed scatters, the second



order matrices are used. The second order matrices are derived from the scattering matrix. In case of reciprocity condition in which  $S_{HV} = S_{VH}$  then the vectorized format of the scattering matrix is given in form of lexicographic basis and Pauli basis [12].

In case of Pauli format:

$$K_P = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix} \quad (2)$$

where  $K_P$  is the Pauli vector. By multiplying this vector with its complex conjugate transpose the coherency matrix  $T_3 = K_P K_P^*$  is obtained:

$$T_3 = \begin{bmatrix} \langle S_{HH} + S_{VV} \rangle^2 & \langle S_{HH} + S_{VV} \rangle \langle S_{HH} - S_{VV} \rangle^* & 2\langle S_{HH} + S_{VV} \rangle \langle S_{HV} \rangle^* \\ \langle (S_{HH} - S_{VV})(S_{HH} + S_{VV})^* \rangle & \langle S_{HH} - S_{VV} \rangle^2 & 2\langle (S_{HH} - S_{VV}) \langle S_{HV} \rangle^* \rangle \\ 2\langle S_{HV} \rangle \langle S_{HH} + S_{VV} \rangle^* & 2\langle S_{HV} \rangle \langle S_{HH} - S_{VV} \rangle^* & 4\langle S_{HV} \rangle^2 \end{bmatrix} \quad (3)$$

### 3.3 Speckle Filter

Speckle appearance in radar images is caused by the coherent interference of waves reflected from many elementary scatters [13]. Speckle can be reduced using multi-look observations, which can be achieved during the image construction, or a speckle-reduction filter performed by the user. In order to achieve optimal speckle reduction in imagery refined lee filter was used. It was used since it has proven to be good in preserving polarimetric information for distributed targets. We tested different filter sizes and the best results were achieved by using a  $7 \times 7$  filter. It is based on statistical correlation between channels without introducing cross talk [14], [15].

### 3.4 Geometric Correction

The most significant step in SAR data pre-processing is the geometric correction. The original measure of SAR system is the slant range so, the image is recorded in slant range system [13]. With the slant range the image can't be visually interpreted because each pixel is compressed and can't also be display with the correct size. So, it needs to be converted into ground range. Geometric correction transfers the slant range image to ground range. The digital elevation model (DEM) of the study area was applied during performing the geometric correction. The study area is generally flat so, the terrain effects of layover and shadowing are neglected. MapReady tool has been used in conducting this part. It is a Remote Sensing

Tool kit developed by Alaska Satellite Facility and embedded in the PolSARpro software.

## 3.5 Feature Extraction

The main objective of the decomposition of the matrix representation (e.g. coherency or covariance matrices) is extracting parameters that carry information about the structural and compositional contents of the ground target or land cover from the measured backscatter. The matrix can be first order (e.g. the scattering matrix in equation 1) or second order (e.g. the coherency matrix in equation 3). The most commonly-used decomposition is that of the scattering matrix which is known as Pauli decomposition. It decomposes the scattering matrix for mono-static case into three components for studying the surface properties which are represented as single-bounce ( $S_{HH} + S_{VV}$ ), double-bounce ( $S_{HH} - S_{VV}$ ), and volumetric ( $S_{HV}$ ) scattering mechanisms [16], [17]. These components are represented by the Pauli vector as in equation 2.

### 3.5.1 Pauli decomposition

The Pauli decomposition parameters are the elements included in the vector of equation 4. The first, second and third element in the vector represent the single-bounce, double-bounce and volume scattering, respectively. This decomposition is the most common and more appropriate for coherent targets (with identifiable structures) compared to other coherent decomposition methods. The Pauli decomposition is the most effective and useful for exposing natural targets, but not ideal for highlighting man-made targets [18]. The scattering matrix  $[S]$  can be written as:

$$[S] = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} = \alpha \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \beta \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \gamma \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (4)$$

Where  $\alpha = (S_{HH} + S_{VV})/\sqrt{2}$ ,  $\beta = (S_{HH} - S_{VV})/\sqrt{2}$  and  $\gamma = \sqrt{2} S_{HV}$ .

Using Pauli decomposition, often  $\alpha$ ,  $\beta$ , and  $\gamma$  components are represented as blue, red, and green respectively for visual interpretation. The polarization color composite of the used image in Pauli basis is displayed in figure 5.

### 3.5.2 Polarimetric signature

The polarimetric signature describes the scattering coefficient as a function of any assumed transmit and receive antenna polarization states (linear, circular,



and elliptical). It allows measure the variation of the scattering coefficient with polarization so that different targets show different polarization signatures [19]. Although many targets can produce similar plots, the shape of the plots as well as the pedestal height, provide clues about the type of scattering dominant from the target. The polarimetric signatures are very sensitive to the orientation of the target relative to the radar line of sight [20]. The angle of the semi-major axis with the horizontal axis (x-axis) is the orientation angle ( $\psi^\circ$ ) ranging from  $0^\circ$  to  $180^\circ$ . Ellipticity defines the oval shape of the ellipse shown as ( $\chi^\circ$ ) as depicted in figure 4. Linear polarizations have an ellipticity angle of  $0^\circ$ , while circular polarizations have ellipticity angles of  $45^\circ$ . Although all orientations are represented in the plot, the commonly used linear polarizations have orientation angles of  $0^\circ$  (H) or  $90^\circ$  (V) [3]. The polarization signature  $\sigma^0$  is represented by the following equation 5:

$$\sigma^0 = K J_r^T \langle M_s \rangle J_t = \sigma^0(\chi_r, \psi_r, \chi_t, \psi_t) \quad (5)$$

where K is constant.  $J_r$  and  $J_t$  are the Stokes vectors at the receiver the transmitter, respectively. The  $\chi$  and  $\psi$  denote the ellipticity and the orientation angles of the electric field vector.

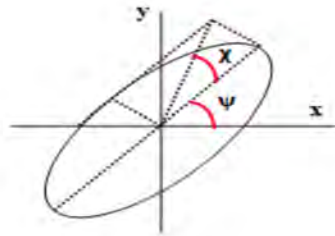


Figure 4: Definition of polarization signature

With the help of the geological map and the ETM-8 image, we choose areas within the Radarsat-2 image for different types of terrain, as illustrated in figure 5. The polarization signature for each resolution element (pixel) represents the sum of the polarization signatures of each object in this pixel. The signatures can be computed on a pixel basis or average over a region that captures the particular feature. Single pixel basis was used for the generation of the various features polarization signatures ((a) Wadi deposits, (b) Chlorite schist, (c) Tonalite, and (d) Radar penetrated areas). Co-polarization and Cross-polarization signature plots were extracted for these pixels.

The following figures (6, 7, 8 and 9) display the calculated polarimetric signatures (backscattered power) of these major features normalized to the

intensity range 0-1 in lin format (Mesh representation).

#### 4. Results and Discussions

There are a wide range of SAR parameters that can be extracted from PolSAR data. Polarimetric signature introduces new concepts for different targets identification. de Mattheaïs et al. [21] describe the co-polarization plots as containing the imprinting by the various scattering mechanisms (surface, double-bounce, and multiple/volume). It is clear from the following figures that, the calculated polarization signatures are more or less different from each other. The cross-polarized response behaves in an exactly opposite manner to the co-polarization response. The relative similarity between the two classes Chlorite schist and Tonalite was predictable due to the same physical properties and the influence of surface roughness. Their co-polarization and cross-polarization signatures show some peaks corresponding to the maximum backscattered power. Their pedestal height were (0.52, 0.4) respectively, indicating moderate amount of randomly oriented backscatter (i.e. rough surface scattering mechanisms). It is also clear that, Wadi deposits and Radar penetrated areas have the same behavior in changing polarization bases. In the sense, for both co-polarization and cross-polarization signatures, there is not much variation in the backscattered power response ranging over various ellipticity and orientation angles. Also there is low amount of randomly oriented backscatter (i.e. minimal depolarization) that can be viewed from their low pedestal heights ( $\approx .07$ ), indicating that surface scattering is dominant. These features are not rough enough to cause multiple or volume scattering. Therefore, this result confirms that surface scattering is typical of areas that appear smooth relative to the Radar wavelength.

#### 5. Conclusion

The polarimetric signatures of various features in Radarsat-2 PolSAR image of part of Halayib area (Egypt) were generated and studied. Several types of scattering are usually present within distributed features, although these features often have a dominant scattering mechanism. The polarization signature plots clearly differentiated these features based on their type of scattering. Analysis results demonstrated that rough surfaces like Tonalite, Chlorite schist cause greater multiple scattering as compared to the smoother surfaces like Wadi deposits and Radar penetrated areas. Shape of the co-polarization, cross-polarization signature plots and the pedestal heights indicate and provide information on the dominant scattering mechanism. The pedestal height was also unique for each of the features.





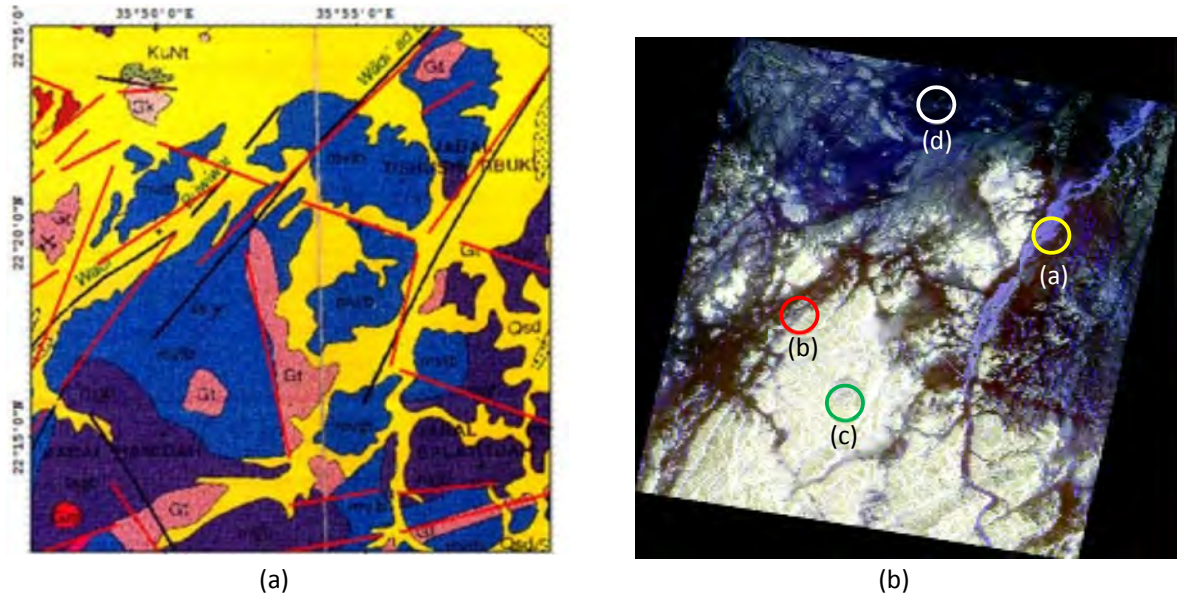


Figure 5: (a) Geological map of the study area, (b) Radarsat-2 PolSAR image in Pauli basis

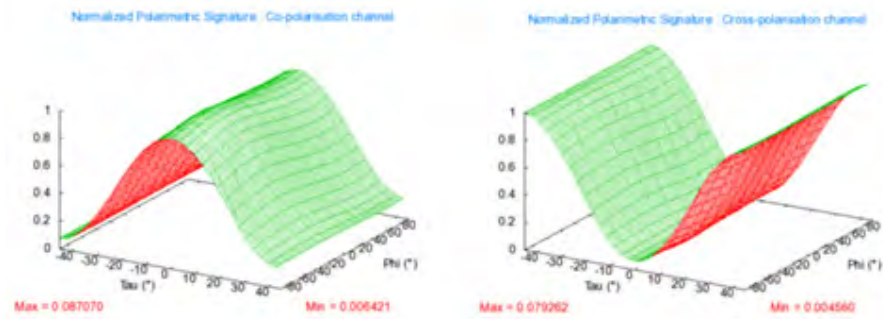


Figure 6: (a) Wadi deposits polarimetric signature

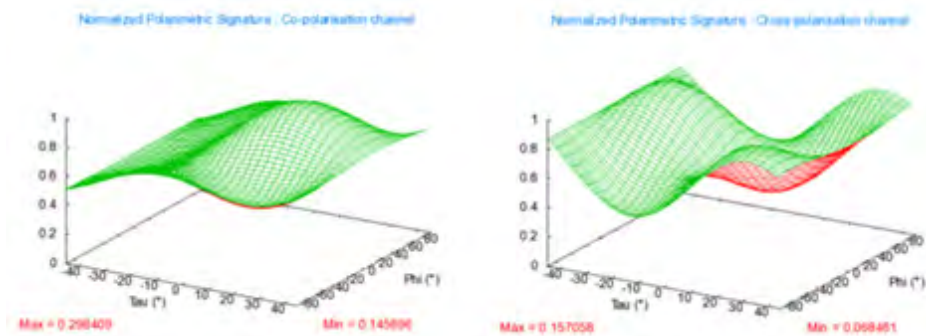


Figure 7: (b) Chlorite schist polarimetric signature



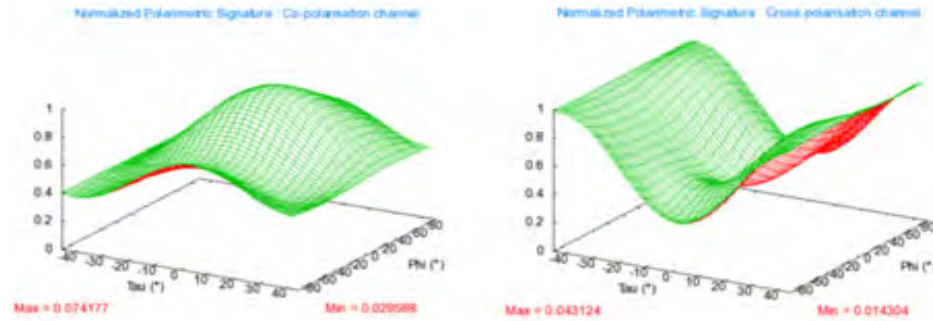


Figure 8: (c) Tonalite polarimetric signature

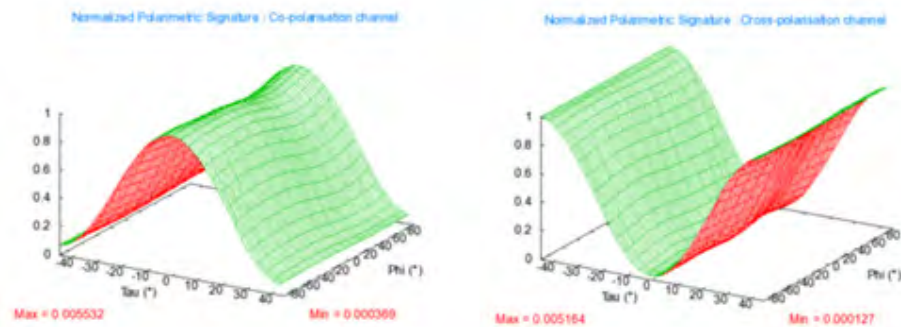


Figure 9: (d) Radar penetrated areas polarimetric signature

Smooth surfaces have low backscattered power values. Rougher surfaces have more multiple scattering and higher pedestal heights. This confirms the sensitivity of pedestal height to multiple and volume scattering. The results illustrate that polarimetric signatures of various features can be an effective criteria for analyzing the different features.

## 6. Acknowledgment

The authors would like to thank the Canadian Space Agency for providing the image used in this study within the framework of their Science and Operational Applications Research – Africa (SOAR-AF) under the agreement of the project (ID: 5128). Also, Special thanks are extended to the National Authority for Remote Sensing and Space Sciences (NARSS) in Egypt for supporting this study.

## 7. References

- [1] F. T., Ulaby, C., Elachi, "Radar Polarimetry for Geoscience Applications", 1990. Artech House Inc., Dedham.
- [2] H. A., Zebker, J. J., van Zyl, "Imaging radar polarimetry: A review", 1991. Proc. IEEE 79, pp. 1583–1606.
- [3] H., McNairn, C., Duguay, B., Brisco, and T.J., Pultz, "The effect of soil and crop residue characteristics on polarimetric radar response," 2001. Remote Sens. Environ., vol. 80, pp.308-320.
- [4] M., Jafari, Y., Maghsoudi, Z. M., Valadan, "Analyzing Polarimetric Signature for Different Features in Polarimetric SAR Data", 2014. IGARSS, IEEE, pp. 2782-2785.
- [5] J. J., Van Zyl, H. A., Zebker, and C., Elachi, "Imaging radar polarization signatures: Theory and observation", 1987. Radio Science, vol. 22, pp. 529-543.
- [6] S. L., Durden, J. J., Van Zyl, and H. A., Zebker, "Modeling and observation of the radar polarization signature of forested areas", 1989. IEEE Transactions on Geoscience and Remote Sensing, vol. 27, pp. 290-301.
- [7] G., De Grandi, J. S., Lee, D., Schuler, and E., Nezry, "Texture and speckle statistics in polarimetric SAR synthesized images", 2003. IEEE Transactions on Geoscience and Remote Sensing, vol. 41, pp. 2070-2088.
- [8] F., Nunziata, M., Migliaccio, and A., Gambardella, "Pedestal height for sea oil slick observation", 2011. IET radar, sonar & navigation, vol. 5, pp. 103-110.



- [9] J. J., Van Zyl, "Unsupervised classification of scattering behavior using radar polarimetry data", 1989. IEEE Transactions on Geoscience and Remote Sensing, vol. 27, pp. 36–45.
- [10] T. W., Ray, T. G., Farr, and van J. J., Zyl, "Polarization signatures for abandoned agricultural fields in the Manix Basin area of the Mojave Desert: can polarimetric SAR detect desertification?", 1992. In: Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS '92), Houston, pp. 947–949.
- [11] J. S., Lee, M. R., Grunes, and E., Pottier, "Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR", 2001. IEEE Transactions on Geoscience and Remote Sensing, vol. 39, pp. 2343–2351.
- [12] V., Nandan, "Retrieval of volume scatterer information using three-components decomposition model, based on compact polarimetry", 2012. M.Sc. Thesis. Reading University: Twente.
- [13] J. S., Lee, and E., Pottier, "Polarimetric radar imaging: from basics to applications", 2009. Taylor & Francis Group, New York.
- [14] C., Niu, G., Zhang, J., Zhu, S., Liu, and D., Ma, "Correlation coefficients between polarization signatures for evaluating polarimetric information preservation," 2011. IEEE Geosci. Remote Sens. Lett., vol. 8, pp. 1016–1020.
- [15] M., Salehi, M. R., Sahebi, and Y., Maghsoudi, "Improving the accuracy of urban land cover classification using radarsat-2 polSAR data", 2013. IEEE journal in selected topics in applied earth observations and remote sensing. vol. 7, pp. 1394 – 1401.
- [16] J. R., Huynen, "Measurement of the Target scattering matrix", 1965. Proceedings of the IEEE. vol. 53, pp. 936-946.
- [17] D. H., William, "Exploration of data fusion between polarimetric radar and multispectral image data", 2012. MSE Thesis. Reading University: California State University.
- [18] L., Zhang, J., Zhang, B., Zou, and Y., Zhang, "Comparison of methods for target detection and applications using polarimetric SAR image", 2008. Progress in Electromagnetic Research, PERS Online, vol. 4, pp. 140–145.
- [19] K., Arai, "Polarimetric SAR Image Classification with High Frequency Component Derived from Wavelet Multi Resolution Analysis: MRA", 2011. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol.2, No. 9.
- [20] R. Z., Schneider, K., Papathanassiou, I., Hajnsek, and A., Moreira, "Polarimetric interferometry over urban areas: Information extraction using coherent scatterers", 2005. In Proc. IEEE IGARSS, Seoul, Korea, Jul. 25–29, pp. 1089–1092.
- [21] P., de Matthaeis, G., Schiavon, and D., Solimini, "Effect of scattering mechanisms on polarimetric features of crops and trees", 1994. International Journal of Remote Sensing, vol. 15, pp. 2917–2930.

## Biographies



Ayman H. Nasr is an Associate Research Professor in National Authority for Remote Sensing and Space Sciences. He is currently the Head of the Data Reception, Analysis and Receiving Station Affairs Division. He received his B.Sc., M.Sc. and Ph.D. in

Electronics and Communications from Faculty of Engineering, Cairo University. His research expertise includes image processing and GIS in different fields of Remote Sensing. He has published 45 scientific papers in various national and international academic journals and conference proceedings. He has contributed in publishing 4 atlases and participated in more than 50 projects and researches related to remote sensing fields. He supervises and graduated several Masters and PhD students. Dr. Nasr is currently a member of the associate editors of the Egyptian Journal of Remote Sensing and Space Sciences. He is also a member of the Egyptian Society of Remote Sensing, GIS and Space sciences. Mail: [aymanasr@hotmail.com](mailto:aymanasr@hotmail.com).



Hind Z. Abdelhamid is an engineer in National Authority for Remote Sensing and Space Sciences. She is currently data analyzer at the Data Reception, Analysis and Receiving Station Affairs Division. She received her B.Sc. and M.Sc. in Computers and Systems Engineering Department, Al-Azhar University. Her area of interest includes image processing and GIS in different fields of Remote Sensing. Mail: [sar.hend87@gmail.com](mailto:sar.hend87@gmail.com)







## Fast Content Based Image Retrieval Using Zernike Moments

Nidhi Goel<sup>1</sup> and Ekta Walia<sup>2</sup>

- 1) Assistant Professor, Computer Science, Maharishi Markandeshwar University  
Mullana, Ambala, 133203, Haryana (India)
  - 2) Associate Professor, Department of Computer Science, South Asian University  
New Delhi 110067 (India)
- Telephone: +91 (0) 11 2671 1471, goelnidhi8@gmail.com, wekta@yahoo.com

### Abstract

Content Based Image Retrieval (CBIR) is an important research area for manipulating large amount of image databases and archives. In a broad sense, features include visual features like color, texture, shape etc. In order to extract features of an image, various feature extraction methods are available. One of them is moment description. The Zernike Moment Descriptor is a moment based Shape Descriptor. It has many desirable properties such as rotation invariance, robustness to noise, expression efficiency and fast computation for describing the shapes of patterns. In this paper, we perform fast Content Based Image Retrieval (CBIR) of images from a database for the given query image. We have shown how fast computation of radial polynomials for computing Zernike Moments (ZMs) leads to the fast retrieval of relevant images according to the similarity measure calculated between features of the query image and images of the image database.

**Keywords:** Content Based Image Retrieval; Zernike Moments; Fast Computation.

### 1. Introduction

An image retrieval system is a computer system for browsing, searching and retrieving images in an image database. Text based and content based are the two techniques for search and retrieval in image database. Text based retrieval is non-standardized because different users use different keywords. Text descriptions are sometimes subjective and incomplete because it cannot depict complicated image features very well. Content based image retrieval (CBIR) technique use image content to search and retrieve images. Content based image retrieval system was introduced to address the problems associated with text-based image retrieval. It is based on extracting and comparing the visual attributes of the images. Examples of visual attributes are color,

texture, shape, and motion parameters. The users usually formulate query image and present to the system. The system extracts the visual attributes of the query image in the same mode as it does for each database image, and then identifies images in the database whose feature vectors match those of the query image, and sorts the best similar objects according to their similarity measure.

An elective shape descriptor is a key component of multimedia content description, since shape is a fundamental property of an object. There are two types of shape descriptors: contour based shape descriptors and region based shape descriptors [1]. Contour based shape descriptors may not be suitable for complex shapes that consist of several disjoint regions such as trademarks or logos, emblems, clipart and characters, including various shapes in natural scenes. Region based shape descriptors, such as moments, are more reliable for shapes that have complex boundaries, because they rely not only on the contour pixels but also on all pixels constituting the shapes [1]. The Zernike Moment Descriptor is a region based shape descriptor.

The drawback of regular moments is that there is redundant information in the moments since the bases are not orthogonal and high-order moments are sensitive to noise [2]. In order to retrieve an image from a large database the descriptor should have enough discriminating power and immunity to noise. In addition, the descriptor should be invariant to scale and rotation.

The Zernike moment descriptor has such desirable properties: rotation invariance, robustness to noise, expression efficiency and fast computation for describing the various shapes of patterns [3]. With a proper normalization method, scale invariance can also be achieved [4].

There are many problems in the existing CBIR systems and are given as below:

- **More Computation Time:** Time taken to retrieve the relevant images is more in the current CBIR systems.



- **More Retrieval Time:** Images should be retrieved in a specified time but time taken to retrieve the relevant images by current CBIR systems is too much.
- **Accuracy:** In most of the CBIR systems relevant images are not retrieved on the top ranked results.

In this paper, we have overcome all these problems by making use of fast q-recursive method to calculate radial polynomials of Zernike Moments [11]. The proposed method gives best performance in terms of total retrieval time and accuracy.

The remainder of the paper is organized as follows: Section (2) focuses on literature. Section (3) emphasizes on our proposed work. In Section (4) we have discussed the results. Final Section (5) is the conclusion and future work.

## 2. Literature Review

Problems with traditional methods of image indexing have led to the rise of interest in techniques for retrieving images on the basis of automatically-derived features such as color, texture and shape – a technology now generally referred to as Content Based Image Retrieval. Lu et. al [5] have used the ‘linear scaling to unit variance’ normalization method to equalize each dimension of the feature vector. A fast search method named equal-average K nearest neighbor search is then used to find the first K nearest neighbors of the query feature vector as soon as possible based on the squared Euclidean distortion measure. Experimental results show that the proposed retrieval method can largely speed up the retrieval process, especially for large database and high feature vector dimension [5].

A new approach to content based image retrieval by texture solved three problems: high computational time, handling high dimension data, and comparing images consistent with human perception. To decrease the computational time, a new strategy was presented to extract an image feature with high retrieval accuracy [6]. The method fast access by content is based on the image tree of contours defined for graphic images, as well as on the one dimensional complex Fourier transform of the contours. In this way, problems connected with the image normalization concerning translation, rotation, scaling, reflection and intensity are currently solved. The most informative data of the image are ordered by importance in a key of fixed length, on which the fast access is performed using the well-known index access methods of a conventional database management system (DBMS). The method is tested on a database of about 4000 images of hallmarks [7].

A novel content based image retrieval data structure is developed [8]. It can improve the searching efficiency significantly. All images are organized into a tree, in which every node is comprised of images with similar features. Images in a children node have more similarity (less variance) within themselves in relative to its parent. Upon the addition of new images, the tree structure is capable of dynamic ally changing to ensure the minimization of total variance of the tree. Subsequently, a heuristic method has been designed to retrieve the information from this tree.



Recently wavelets are used along with ZMs [9]. Dual Tree Complex Wavelets and Fourier Features are extracted for whole image database and stored separately. Afterwards query image is given and system retrieves images only on the basis of features of ZMs as it is a region based descriptor. The irrelevant images are thus filtered out and the selected database is then compared with query image on the basis of Dual Tree Complex Wavelets and Fourier Features.

A fast and robust color feature extraction method for effective content based image retrieval is used [10]. In color feature extraction, since cylindrical HSV color space is not perceptually uniform, the color quantization and similarity measure method based on a cylinder-cone transform is improved. A new rectangular approximate image segmentation is used. A significance function is also used to reflect the importance of different position in image, and improve the segmentation and retrieval performance.

## 3. Proposed Work

The way to increase the responding ability is to speed up the procedure of content retrieval. Fast methods of visual feature extraction can be used to make the process of content based image retrieval fast. It is hard for the traditional systems to increase the responding ability. Chong et. al [11] carry out an extensive survey of fast methods and propose a new method which is popularly known as q-recursive method, where  $q$  represents the repetition term in ZMs. The q-recursive method is reported to be the best and fastest method among all the recursive methods to compute Zernike polynomials [12].

### 3.1 Objectives of the proposed work:

- The main objective is to make CBIR system fast by using the fast algorithm to compute Zernike Moments. Here we take image feature (ZMs) as an index to that image and retrieve the relevant images.
- We implement the fast CBIR system which takes into consideration the low level features of image which are more comprehensive when compared to high level features.

In this paper, we use an efficient and fast algorithm for the recursive computation of Zernike polynomials. The CBIR process is made faster by using q-recursive method for fast access of contents from database of images. Fast CBIR is done in terms of:

- **Computation Time:** The computation of retrieval algorithm should not take time beyond certain prescribed limit.
- **Retrieval Time:** Images should be retrieved in a specified real time.
- **Accuracy:** Required relevant images should be retrieved on the top of the list of ranked images.

### 3.2 Architecture of Fast CBIR using q-recursive Method:

The architecture of fast CBIR system is shown in Figure 3.2. There are two issues in building this Fast CBIR system:



- 1) Every image in the image database is to be represented efficiently by extracting significant features using ZMs.
- 2) Relevant images are to be retrieved using similarity measure between query image and database images.

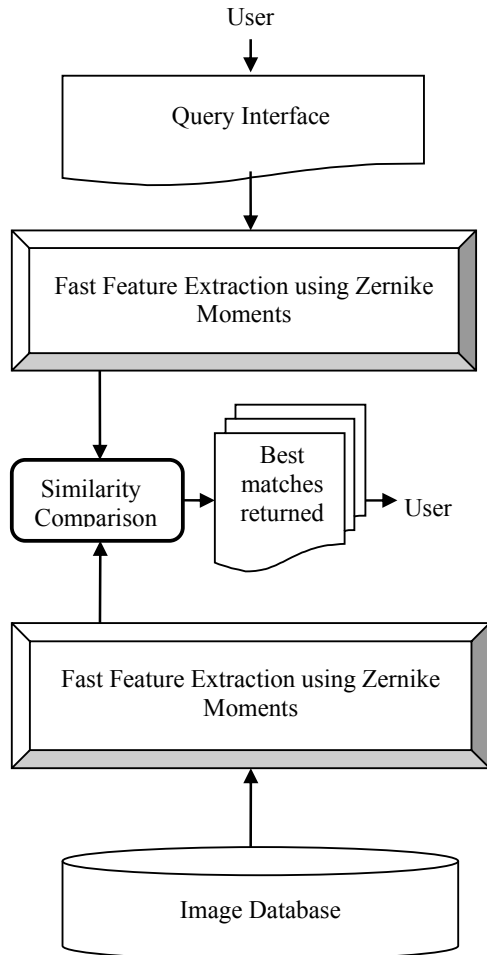


Fig 3.2: Architecture of Fast Content Based Image Retrieval System

To achieve this functionality, the fast CBIR has two main components: database population (the process of creating an image database) and a query. During the population, images are processed to extract features describing their shapes and the shape features are stored in a database. In the query phase, the user composes a query image. Features are generated from the query and then input to a matching engine that finds images from the database with similar features.

The performance of the fast CBIR system can be tested by retrieving the desired number of images from the database. The total retrieval time taken is the main performance measure in the method used for fast CBIR system. The average retrieval rate is known as the average percentage number of images belonging to the same image as the query image in the top 'N' matches. 'N' indicates the number of retrieved images.

### 3.3 Algorithmic Steps for Fast Content Based Image Retrieval

The q-recursive method is used to make the process of CBIR fast. The traditional approach to compute ZMs known as direct radial polynomial formulation is also implemented. It is worth mentioning that image is inscribed within a circle for computing ZMs with q-recursive method. In case of traditional ZMs, they are however computed by taking a unit disk inscribed in the image.

#### Steps for Content Based Image Retrieval for both q-recursive and direct method:

1. The user gives an image  $I$  as the query image.
2. a) Get the ZMs of the query image.  
b) Find the time taken to calculate ZMs of the query image.
3. Find the Mean Square Reconstruction Error ( $MSRE$ ) for the query image.
4. a) For image  $i$  in the database obtain the ZMs.  
b) Find the time taken to calculate ZMs of the image  $i$ .
5. Find the  $MSRE$  for image  $i$  in the database.
6. Calculate the Euclidean distance between the two sets of ZMs and store them in an array.
7. Increment  $i$ . Repeat from step 4.
8. For each image  $i$  in the database, if  $MSRE < 1.0$  sort the array of Euclidean distances.
9. Calculate the total time taken for CBIR.
10. Retrieve the top 'N' results; N is the number of top ranked images.
11. Mark the top ranked retrieved images as relevant or irrelevant.
12. Obtain the Precision and Recall metrics for performance evaluation.

Using the above algorithm, the most relevant images are searched for in the image database. The Euclidean distance is calculated between the query image and every image of the database. This process is repeated until all the images in the database have been compared with the query image. Upon completion of the calculation of Euclidean distances, we have an array of Euclidean distances, which is then sorted. The 'N' topmost images are then displayed as a result of the search and then marked as relevant or irrelevant.

## 4. Results and Discussion

### 4.1 Dataset and Image Features

We performed our experiments over an image collection. This collection contains around 25 images associated with textual description and cataloged into 4 broad categories: butterfly, food, medical images and labels. Image features used in our experiments are Zernike Moments. q-recursive method is used for the feature extraction using Zernike Moments. Euclidean distance function is used for the similarity measure. Figure 4.1, Figure 4.2, Figure 4.3 and Figure 4.4 show the four categories of images used as datasets.





Figure 4.1 Class 1: Butterfly Images



Figure 4.2 Class 2: Food Images



Figure 4.3 Class 3: Medical Images



Figure 4.4 Class 4: Labels Images

## 4.2 Evaluation Method

The retrieval performance of any CBIR system is inherently limited by the quality of the features used to represent images. In this paper, we have performed a precision and recall study to evaluate the performance of both the methods i.e. traditional ZMs computation and fast recursive ZMs computation. This paper focuses on the ability of our relevance feedback mechanism to converge to a desired result. We run an initial query with images chosen at random from the collection and save the ranked result as the relevant list and then these results are refined.

For a given query image, mean-square reconstruction error (*MSRE*) [12] is denoted by  $\mathcal{E}_L$  and is calculated as:

$$\mathcal{E}_L = \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \{f(x_i, y_j) - \hat{f}(x_i, y_j)\}^2}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x_i, y_j)} \quad (1)$$

Here,  $f$  is the original image and  $\hat{f}$  is the image reconstructed from ZMs. Euclidean Distance is measured for the images having *MSRE* less than 1.0. The system first retrieves a list of ranked images according to the Euclidean Distance as a similarity measure. Then, the user marks the retrieved images as relevant (positive examples) to the query or not relevant (negative examples). The system will refine the retrieval results based on the feedback and present a new list of images to the user.

### 4.2.1 Euclidean Distance

Euclidean distance is a very commonly used distance measurement. Euclidean distance actually measure dissimilarity. Small distance means small dissimilarity and large similarity. Euclidean distance is basically the sum of squared distances between two vector values and obtained by the following equation:

$$D(Q, I) = \sqrt{\sum_{i=1}^N (f_j^Q - f_j^I)^2} \quad (2)$$

### 4.2.2 Metrics used for Performance Evaluation

Precision and recall are two widely used metrics for evaluating the correctness of a pattern recognition algorithm. For a given query  $q$ , the data set of images in the database that are relevant to the query  $q$  is denoted as  $R(q)$ , and the retrieval result of the query  $q$  is denoted as  $Q(q)$ .  $Q(q)$  is the total number of both irrelevant and relevant records retrieved.

- **Recall (R):** The recall is the fraction of relevant images that is returned by the query. It is the ratio of the number of relevant records retrieved to the total number of relevant records in the database.

$$\text{Recall} = \frac{R(q)}{\hat{R}(q)} \quad (3)$$

where  $\hat{R}(q)$  is the total number of relevant records in database i.e. in the same category to which the query image belongs.

- **Precision (P):** The precision of the retrieval is defined as the fraction of the retrieved images that are indeed relevant for the query. It is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved.

$$\text{Precision} = \frac{R(q)}{Q(q)} \quad (4)$$

### 4.4 Experimental Study

The q-recursive method is the fastest to compute ZMs [12]. The CBIR is made fast by using the q-recursive method. The direct method to compute ZMs [13] is also implemented in CBIR. The two approaches are compared and analyzed for their accuracy, total retrieval time and numerical stability.



Table 1: Recall and precision for Class 1(Butterfly) Image Size: 64×64

Moment Order	Method used	No of relevant images	Recall and Precision Values (in %age)
5	Direct	2	R=40,P=13.33
	q-recursive	4	R=80,P=26.66
10	Direct	2	R=40,P=13.33
	q-recursive	4	R=80,P=26.66
20	Direct	2	R=40,P=13.33
	q-recursive	4	R=80,P=26.66
30	Direct	2	R=40,P=13.33
	q-recursive	4	R=80,P=26.66
35	Direct	0	R=0,P=0
	q-recursive	4	R=80,P=26.66

Table 2: Recall and precision for Class 2(Food) Image Size: 64×64

Moment Order	Method used	No of relevant images	Recall and Precision Values (in %age)
5	Direct	2	R=50,P=13.33
	q-recursive	3	R=75,P=20
10	Direct	2	R=50,P=13.33
	q-recursive	3	R=75,P=20
20	Direct	1	R=25,P=6.66
	q-recursive	3	R=75,P=20
30	Direct	1	R=25,P=6.66
	q-recursive	3	R=75,P=20
35	Direct	0	R=0,P=0
	q-recursive	3	R=75,P=20

Table 3: Recall and precision for Class 3(Medical) Image Size: 64×64

Moment Order	Method used	No of relevant images	Recall and Precision Values (in %age)
5	Direct	0	R=0,P=0
	q-recursive	0	R=0,P=0
10	Direct	0	R=0,P=0
	q-recursive	2	R=40,P=13.33
20	Direct	0	R=0,P=0
	q-recursive	3	R=60,P=20
30	Direct	0	R=0,P=0
	q-recursive	3	R=60,P=20
35	Direct	0	R=0,P=0
	q-recursive	3	R=60,P=20

Table 4: Recall and precision for Class 4(Labels) Image Size: 64×64

Moment Order	Method used	No of relevant images	Recall and Precision Values (in %age)
5	Direct	3	R=42.85,P=20
	q-recursive	5	R=71.42,P=33.33
10	Direct	2	R=26.57,P=20
	q-recursive	5	R=71.42,P=33.33
20	Direct	2	R=26.57,P=13.33
	q-recursive	5	R=71.42,P=33.33
30	Direct	2	R=26.57,P=13.33
	q-recursive	5	R=71.42,P=33.33
35	Direct	0	R=0,P=0
	q-recursive	5	R=71.42,P=33.33

Table 5: Total Retrieval Time (in seconds) Image Size: 64×64

Moment Order	Total Retrieval Time taken by Direct method (in sec)	Total Retrieval Time taken by q-recursive method (in sec)
5	8.609	0.032
10	27.969	0.126
20	109.627	0.357
30	246.173	0.735
35	-	0.888
50	-	1.639

Table 6:  $MSRE(\mathcal{E}_L)$  for Class 1 Image Size: 64×64

Moment Order	$MSRE$ Direct method	$MSRE$ q-recursive method
5	0.494267	0.091724
10	0.534892	0.058560
20	0.335957	0.061277
30	0.324148	0.051857
35	1.000000	0.026130
50	1.000000	0.020648





- **Numerical Stability:** The numerical stability of both the methods are analyzed in terms of Mean Square Reconstruction Error ( $\varepsilon_L$ ) [12] with respect to the order of moments for a given image size. *MSRE* for an image of Class 1(Butterfly) is shown in Table 6. It shows that *MSRE* for the direct method becomes 1 for moment order 35 while for q-recursive method it keeps on decreasing with the moment order. It shows as the moment order increases, numerical stability of direct method becomes very poor. Numerical stability of q-recursive method increases with respect to the order of moments. So q-recursive method is more stable [12] and retrieves maximum number of relevant images.
- **Total Retrieval Time:** The total time taken by both the methods to retrieve the relevant images for CBIR is shown in Table 5. For the moment order 30 total time taken by direct method is 246.173 sec and by q-recursive is 0.735 sec. It shows that the proposed method is many times faster than the direct ZMs based CBIR. Thus the proposed method makes the CBIR system fast by retrieving maximum number of relevant images in minimum amount of time.
- **Accuracy:** Accuracy is calculated on the basis whether the retrieved images are relevant or irrelevant. It is measured in terms of Recall (R) and precision (P). Tables 1, 2, 3 and 4 show the R and P for all class (i.e. 1 to 4) images respectively. The tables show that R and P for direct method decreases with the moment order. In the case of q-recursive method either it increases with the moment order or remains constant. In Table 4, R and P for q-recursive method increases with moment order while in table 1 it remains constant. Recall value in most of the cases is more than 75%. It implies that the system retrieves the maximum number of relevant images from database using q-recursive method.

Hence, q-recursive method tends to give consistent performance whereas direct method (probably due to inherent instability) performs inconsistently and is relatively less accurate.

## 5. Conclusions and Future Work

### 5.1 Conclusions

The CBIR is made fast by using the fastest method to compute Zernike Moments. There are two approaches to compute ZMs of an image function—direct radial polynomial formulation and q-recursive method. The two approaches are compared and analyzed for their accuracy, time complexity and numerical stability. Based on the above analysis, the following conclusions are drawn.

- The q-recursive method is very fast. Time taken by q-recursive method for CBIR is lowest of all the fast methods available to retrieve the images on the basis of content.
- Direct method used for CBIR is not numerically stable, as the moment order increases, numerical stability becomes very poor. So number of relevant images retrieved becomes less and performance of CBIR degrades.
- The q-recursive method used for radial polynomial formulation of ZMs is numerically stable even for

very large orders of moment. The numerical stability increases with respect to image size. Therefore, it retrieves maximum number of images that are the most relevant.

### 5.2 Future Work

We intend to integrate various other features i.e. color, texture with ZMs features.

## 6. References

- [1] Whoi-Yul Kim, Yong-Sung Kim, "A region-based shape descriptor using Zernike moments" Signal Processing: Image Communication, vol.16, issue 1-2, pp. 95-102, September 2000.
- [2] S. Satini, R. Jain, "Similarity Measures", IEEE Transactions on pattern analysis and machine Intelligence, vol.21, no.9, pp. 871-883, September 1999.
- [3] A. Khotanzad and Y. H. Hongs, "Invariant image recognition by Zernike Moments", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 489-497, 1990.
- [4] V. Gudivada and V. Raghavan, "Content based image retrieval systems" IEEE Computer, vol. 28, no. 9, pp. 18-22, 1995
- [5] Zhe-Ming Lu, Hans Burkhardt and Sebastian Boehmer, "Fast Content-Based Image Retrieval Based on Equal-Average K-Nearest-Neighbor Search Schemes", Advances in Multimedia Information Processing , vol. 4261, pp. 167-174, 2006.
- [6] Mira Park, Jesse S. Jin, Laurence S. Wilson, "Fast Content-Based Image Retrieval Using Quasi-Gabor Filter and Reduction of Image Feature Dimension", Fifth IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 0178, 2002.
- [7] D.T. Dimov, "Fast, Shape Based Image Retrieval", Proceed. of Comp Sys Tech 2003, ACM, pp. 296 – 302, 2003.
- [8] Jinyan Chen, Jizhou Sun, Rongteng Wu and Yaping Zhang, "Fast content based image retrieval using dynamic cluster tree", Proc. SPIE 6833, 68332A, 2007.
- [9] Anjali Goyal, Ekta Walia, Harvinder Singh Saini, "Enhanced Retrieval Accuracy with ZMs using Dual Tree Complex Wavelets and Fourier Features", ICGST-GVIP Journal, vol. 10, no. 3, pp. 27-34, August 2010.
- [10] Chi-Man Pun, Chan-Fong Wong, "Fast and Robust Color Feature Extraction for Content based Image Retrieval" International Journal of Advancements in Computing Technology, vol. 3, no. 6, July 2011.
- [11] Chee-Way Chong, P. Raveendra, R. Mukundan, "A comparative analysis of algorithms for fast computation of Zernike Moments", Pattern Recognition, vol.36, pp. 731 – 742, 2003.
- [12] Chandan Singh and Ekta Walia, "Algorithms for Fast Computation of Zernike Moments and their Numerical Stability", Image and Vision Computing, vol. 29, no. 4, pp. 251-259, 2011.
- [13] Chandan Singh, "Improved quality of reconstructed images using floating point arithmetic for moment calculation" Pattern Recognition, vol. 39, pp. 2047 – 2064, 2006.



## Biographies



**Nidhi Goel** received her B.Tech. Degree in Computer Science from Kurukshetra University, India, in 2007 and M.Tech Degree from Maharishi Markandeshwar University, Ambala, Haryana, India, in 2010. Currently she is working for Maharishi Markandeshwar University at Ambala, Haryana, India. Her academic achievements include University Gold medal in Post Graduation. Her research interests include Image Processing and Database Systems.



**Ekta Walia** received her Bachelor's degree in Computer Science from Kurukshetra University, India, in 1995 and Masters in Computer Applications as well as Ph.D. (Computer Science) from Punjabi University, Patiala, India in 1998 and 2006 respectively. After starting her professional career as a software consultant with DCM DataSystems, New Delhi, India, in 1998, she served as Lecturer and Senior Lecturer in the National Institute of Technical Teachers Training and Research (NITTTR), Chandigarh, India for 07 years, where she was primarily involved in conducting short term courses for teachers of various engineering and polytechnic colleges. She has also acted as resource person for curriculum development workshops, instruction material development and was responsible for conducting research during her service at NITTTR. From 2007 to 2009, she worked as Reader in the Department of Computer Science, Punjabi University, Patiala, India. Later she served as Professor and Head in the Department of Information Technology, M.M.University, Mullana, India from 2009 to 2011. In both these assignments she was actively involved in teaching and research, mentoring as well as infrastructural development. At present, she is serving as Associate Professor in Department of Computer Science, South Asian University, New Delhi, India. Her academic achievements include University Gold medals in Graduation as well as in Post Graduation. She has authored two textbooks. Her research interests include 3D Rendering, Digital Image Watermarking, Content Based Image Retrieval and Face Recognition.









## Automatic Pulmonary Embolism Detection System

Hao-Hung Tsai, Chiun-Li Chin, Yung-Chih Cheng

*Department of Applied Information Sciences of the Chung Shan Medical University, Taichung, Taiwan.*

*Department of medical imaging, Chung Shan Medical University Hospital, Taichung, Taiwan*

*csh232@csh.org.tw, ernestli@mercury.csmu.edu.tw, isethehey@hotmail.com*

### Abstract

Pulmonary embolism (PE) is a blockage of the main artery of the lung or one of its branches by a substance that has travelled from elsewhere in the body through the bloodstream. Most of the traditional PE detection methods need to depend on the professional physician's judgment. Serious PE will lead to dead. Therefore, it is very important for diagnosis PE. In this paper, we develop an automatic pulmonary embolism detection system for relieving doctor's load. It is divided into five parts- preprocessing, finding pulmonary, vessel searching, vessel tracking and evaluation. The experiment results show our system has 83% precision rate in main vessel detection and 82.6% precision rate in branch vessel detection. Finally, our system can relieve doctor's load according to the result of questionnaire analysis.

**Keywords:** CT, ACM, pulmonary embolism

### 1. Introduction

Pulmonary arterial embolism is a common emergency in medical treatment. According to the statistic from European Society of Cardiology, there are 100,000 cases per year in France; there are 65,000 cases per year in England and Wales Borderlands, and there are at least 60,000 cases per year in Italy. If there has an accurate diagnosis and medical treatment opportunely, it effective decreases the death rate to 2% ~ 8%. The general detection types can be separated into four types which are D-miner, X-ray, nuclear medicine and computer tomography (CT) etc. The first type D-II (D-Dimer) is the preparatory method to detect the PE. The positive rate of PaO<sub>2</sub> is 65.9%, PaCO<sub>2</sub> is 53.1%, and the accurate rate of D-Dimer is 70%. Even the sensitive rate is high, the accurate rate is low. X-ray can show infiltration, atelectasis, diaphragmatic eventration and pleural effusion in image, especially in the diagnosis of Hampton based on pleura and the pulmonary artery with Westermark expanded shows the great contributions. But it is inaccurate and also weak to find the position of the embolism. Nuclear medicine uses the behavior of the radioisotope. It offers accurately on diagnosis with many diseases and makes the treatment more efficacious. But the disadvantage of using nuclear medicine is spending

time and there is a problem of registration. Hence, we use CT as the data source to accomplish our system. The purpose of the system we proposed is constructing a system which can provide a high accurate rate helping the doctor's detection, decreasing the workload of doctors and the error occurrence.

In general, the velocity flow of local blood vessels with embolism will lower than other normal region. If getting the pulmonary blood flow distribution in detail, we can estimate the location and the incidence of pulmonary embolism. Qanadi SD, Hajjam ME and Mesurrolle B [1] uses CTA to analyze whether there are emboli in pulmonary artery and vessel or not. The sensitivity and the discrimination rate were 90% and 94%, respectively. Unfortunately CTA does not show the good accurate rate in branch vessel. Qanadi SD, Hajjam ME and Mesurrolle B [2] discovers using lung scan and CTA to detect the pulmonary embolism, the correction rate was only 60%. Nicholas J. Screation and Harvey O. Coxson [3] indicate that the low sensitivity reason can divide into two parts. The first reason is that there are a large number of branch of vessel and the volumes are too slight to be observed. The doctor has to be very careful to diagnosis, and it may make the doctor too tired causing mistakes. Subsequently, every Computed Tomography Angiography (CTA) image can replace the actual data of lung about 2mm or 5mm. It may cause shifted phenomenon by breathing or heart beating. The shifted image makes the area which has lower CT value look like embolism in the branch. Hence, it will increase the error rate of detection embolism.

U. Joseph Schoeph and Nicolaus Holzknecht [4] proposes that it needed to be improved with CTA analysis to diagnose the embolism in branch vessel. According to statistical analysis, the normal perfusion and reduced perfusion of the region where exist embolism could be distinguished with CTA image after injecting the contrast medium [5]. We use the high or low CT value to determine the perfusion with single image or a sequence images. Wildberger and other researchers [6-8] use the image after injecting the contrast medium to observe the difference between the changes of CT value of lung area. It can reflect the difference of perfusion. Users can observe the unusual



area easily. At the same time, there were many computer science scholars develop an automatic algorithm to detect the location of pulmonary embolism [9-13]. They use a lot of image processing technology and artificial intelligence technique such as finding the outline and using neural networks, but the results of their accuracy and precision is not good enough.

Before performing the detection of pulmonary embolism, we need to find out the location of lung region in CTA image. There are some methods to segment each area in image with image processing, such as [14-15]. In addition, there are other methods to find the position of lung by using the features of lung or some symptoms. Park W., Hoffman E.A. and Sonka M [16] use the features of bronchial distributions. Blechschmidt R.A., Werthschützky R and Lorcher U [17] use the features of emphysema. Brown M.S., McNitt-Gray M.F., Mankovich N.J., Goldin J.G., Hiller J., Wilson L.S. and Aberie D.R. [18] uses the fluid air in lung as the pulmonary feature. Dajnowiex, M., Alirezaie, J. and Yongbum Lee, Takeshi Hara, Hiroshi Fujita, Shigeki Itoh and Takeo Ishigaki [19-20] uses pulmonary tuberculum as the feature to find out the position of lung. Besides the above methods which use the features of lung to perform segmentation, some researches, such as [21-23], combine the threshold value of gray image with morphology operators to segment lung area, in which Shojaii, R. Alirezaie and J. Babyn, P. [23] uses the optimal threshold value obtained by automatic calculation to reduce the noise of background and keeping the biggest two areas. In our medical image database, some images have the problem of two lung images superimposition. In order to solve this problem, the calculation has to strengthening. U.J. Schoepf, N. Holzkecht, T.K. Helmberger, et al. [24] uses dynamic programming method to solve the problem of superimposed image. The others studies combined other technologies with knowledge to make the lung region to be separated. For example [16][18], it employs the information of logical structure and provides a powerful segmentation technique. But its shortcoming is high computational complexity. Methods mentioned above simply use the technology of image processing and experiences to get the result for solving the problem. But they donate guarantee a high accuracy, and also do not have a fixed steps to implement.

Subsequently, we discover that the contrast enhancement of local image can help to find the problem area. U.J. Schoepf, N. Holzkecht, T.K. Helmberger, et al [25] proves that using the contrast enhancement of CTA image with vessel part can help the medical diagnosis. C. Zhou, et al. [26] use computer-aided detection (CAD) method to find out the suspected pulmonary embolisms in the image. In visualization, E.A. Chiang, et al. [27] use Maximum Intensity Projection (MIP) method surrounding the heart to search the vessel. It can highlight the higher pixel value region such as the lung image and used it to detect the embolism position. Therefore, it is proper to detect the pulmonary embolism. A.P. Kiraly, D.P. Naidich, and C.L. Novak [28] proposes that the spin projection can segment vascular tree for improving the sensitivity in a few numbers of images. In addition, E. Pichon, C.L. Novak, A.P. Kiraly, D.P. Naidich, A.P. Kiraly, C.L. Novak, D.P. Naidich and Szeliski, R., and Tonnesen, D. and

Terzopoilos D. [29-31] use the material within the vessel to determine the color of the vessel surface in 3D visual aspect. It provides observing all of the blood vessel quickly. These methods we mention above need to perform vessel tracking with 3D navigation method. But 3D navigation in image processing will reduce the efficiency of the system. A.P. Kiraly, D. P. Naidich and C. L. Novak [32-33] propose building the hierarchy tree method. It can raise efficiency of the vessel tracking. Besides, Henri Bouma, Jeroen J. Sonnemans, Anna Vilanova and Frans A. Gerritsen [34] use varied features of embolism such as intensity, shape and size and the location to detect whether there exists embolism in pulmonary artery or not.

According to the above literature review, we discover that several CAD systems for PE have been proposed, but their evaluations have some problems. Firstly, their methods are complicated. Secondly, they use a lot of the range of CT values to remove unnecessary part in the CTA image according to professional doctor's knowledge. Finally, they do not have to process branch vessel in the lung area. For solving these problems, we propose an automatic pulmonary detection system and hope that our system can relieve doctor's fatigue.

The following sections describe details of the proposed method. In section 2, that is our methodologies, we will introduce our proposed detection flowchart and describe each processing procedure in detail. Section 3, it will show the experimental results. And section 4, we will make a conclusion and purpose the research topics in the future.

## 2. Methodologies

First, for observing the state of blood stream in lung, we inject contrast medium into lung. The contrast medium will follow the blood stream and distributes around every vessels in the lung. The lung area which injects contrast medium has higher CT value and looks brighter. The yellow circle area in the Fig. 1 shows the location of embolism. Because blood cannot pass through the embolism area, the contrast medium also cannot pass through it. Hence there does not have the effect of contrast medium, the CT value of embolism is lower than around area where passes through blood with contrast medium. The area which has pulmonary embolism in CTA image shows anomalous dark.



Figure 1 The CTA image with pulmonary embolism

Subsequently, we start to perform the pulmonary embolism detection task. Figure 2 shows the processing procedure of our proposed method. It is divided into four parts: the pre-processing, finding pulmonary, vessel searching and tracking, and using template matching method to evaluate the results between doctor and our system. We will introduce each processing step in detailed.



### 2.1 Pre-processing

The thickness of slice image we used is 1mm. It lets us observing the change of vessel clearly. But, the image has the noise if the thickness of slice image is too thin. The purpose of pre-processing is to obtain an image which can be analyzed. Therefore, we use a smoothing filter to reduce the noise on the original image. The filter uses a  $3 \times 3$  mask to achieve. This mask performs convolution operator with original image from the top of left to bottom of right of the image. The value of central point will be replaced by the mean of the sum of another 8 points. Therefore, we obtain an image which was removed noise.

### 2.2 Finding Pulmonary Location

Next, we have to separate lung area standing alone from background area in the image. Here, we use Active Contour Model (ACM), which is proposed by Michael Kass, Andrew Witkin and Demetri Terzopoulos [35] to obtain the lung area in CTA image. ACM uses the contour points around the target pulmonary convergence method to get the target image contour.

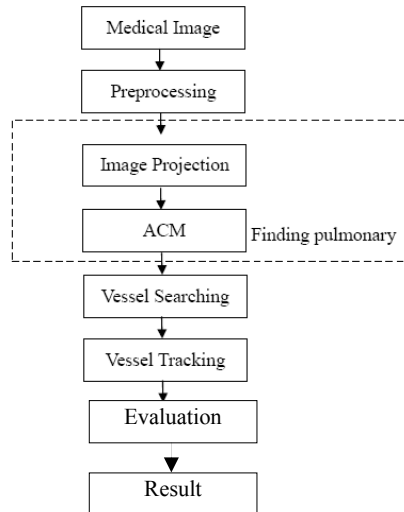


Fig. 2: The flow chart of pulmonary embolism detection

The advantage of using ACM is that it can achieve the goal of finding pulmonary image location automatically. Therefore, we have to set control points around the pulmonary before using ACM. We use image projection method to get those contour points around the target. These control points use B-spline Curve to associate an original contour.

The CT value of body fat is between bone and air. Therefore, when setting the contour points of the target, we can use this feature and combine with the projection on X and Y axis to achieve this task. Figure 3 and Figure 4 show the result of projection on X and Y axis, respectively. We project on X axis and Y axis respectively to get the center of valley. We get the center point of pulmonary in the image with intersection point obtained by the intersecting of center position of X and Y axis.

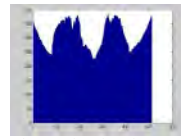


Figure 3 The projection image on X axis

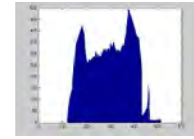


Figure 4 The projection image on Y axis

Here  $(a, b)$  is the center point, and  $a$  is the valley between two obvious peaks on the projection of X axis and  $b$  is the valley between two obvious peaks on the projection of Y axis. Using difference on the projection of X axis can find two huge peaks. The radius  $r$  is calculated by the distance on the first point  $(a_1, b_1)$  of the first huge peak and the end point  $(a_2, b_2)$  of the second huge peak. Hence, based on the characteristic of ACM, we set the contour points around the image of lung. And, the radius is the farther distance calculated from the center point  $(a, b)$  to  $(a_1, b_1)$  and  $(a_2, b_2)$ . The expression equation is as the following.

$$r = \arg \max(\sqrt{(a_1 - a)^2 + (b_1 - b)^2}, \sqrt{(a_2 - a)^2 + (b_2 - b)^2}) \quad (1)$$

When getting a contour circle, we use the angle  $\phi$  obtained by edge detection method to produce a new contour point. We can obtain a contour of fitting pulmonary image edge with iterative step. Figure 5 shows the original image, Fig. 6 shows the initial contour in the image, and Figure 7 shows the image obtained by executing ACM.

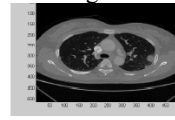


Figure 5 The original CTA

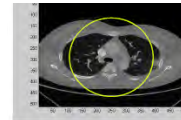


Figure 6 The initialization of ACM

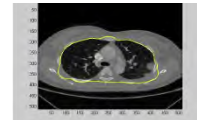


Figure 7 The image after using ACM

### 2.3 Vessel Searching

Figure 8 shows the CTA image before injecting the contrast medium and Figure 9 shows the CTA image with contrast medium. Here, we divide processing step into two parts to perform. One is main vessel processing and the other is branch vessel processing. These two parts do not have the sequence of relationship about executing. The system integrates the results of two parts together and shows the analysis result. The main vessel contains pulmonary artery and aorta. The branch vessel contains other small vessels in lobe. Fig. 10 shows the flow chart of vessel searching.



Figure 8 The CTA image without injecting contrast medium

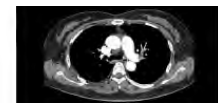


Figure 9 The CTA image after injecting contrast medium

Due to the blood with contrast medium does not flow to the same region at the best time with capture CTA image, the processing of branch vessel and main vessel are different. According to the professional doctor's suggestion, the optimal time to get the appropriate CTA image is the blood with contrast medium just passing through those main vessels.



At this moment, there is higher CT value at the main vessel. About the branch vessel part, the blood with contrast medium does not flow to here yet; the CT value of branch vessel is lower. According to the difference of mentioned above, the methods we used to search main vessel and branch vessel are different. We will introduce the detail from branch vessel searching to main vessel searching.

#### A. Branch vessel searching

First, we use the opening method of morphology of image processing to separate two region of lobe from background. It can help us to perform the following steps more easily. After getting the lobe image, we use histogram equalization method to enhance the contrast of lobe image for observing the detail in lobe.

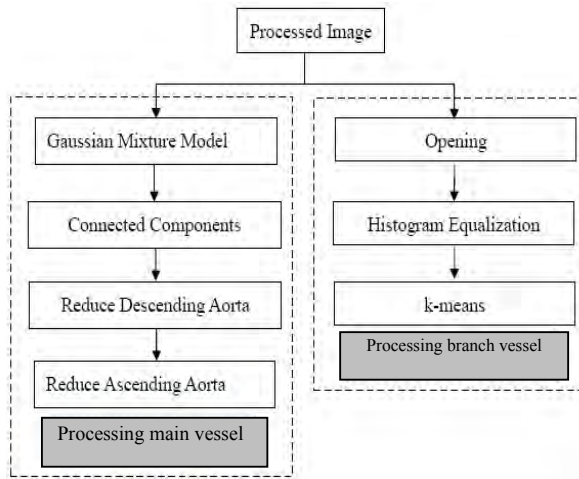


Figure 10: The flow chart of vessel searching

The contrast enhancement can be achieved with eq. 2. And the resulted image are shown on Fig. 11.

$$\text{round}(j) = \lambda \sum n_j \quad \text{where} \quad \lambda = \frac{L-1}{N} \quad (2)$$
  
 $j$  represents gray value,  $\sum n_j$  is the accumulating of  $j$ ,  $L$  is the gray level of the image, and  $N$  is the total pixel number in the image.

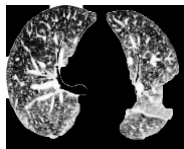


Figure 11 The lobe image after performing histogram equalization

When observing the Figure 11, we discover that the lobe region after performing histogram equalization not only enhances the vessel part clearly, but also enhances other detail parts. For only keeping vessel parts, we use k-means method to keep the more vessel parts for avoiding the interference the later steps. Because the observed data have been injected contrast medium, we will keep the cluster which has the cluster center of highest value. According to the professional doctor's decision and recognition, the data will be divided into 7 clusters. It can satisfy the requirement of detecting pulmonary embolism. Fig. 12 shows the result after performing k-means. Moreover, those areas after histogram equalization and k-

means have very tiny parts for observing. These parts can not recognition by professional doctor. Therefore, we have combined connected components method to remove those very tiny areas in the image. Fig. 13 shows the resulted image after performing connected components.

#### B. Main vessel searching

As we know the pulmonary embolism will occur in the pulmonary vessel, we use Gaussain mixture model (GMM) to separate the main vessel and background. Figure 14 shows the resulted image after performing GMM to analyze the distribution of the pixel intensity in the image.

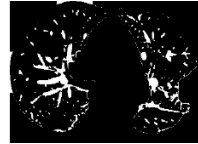


Figure 12 The vessel image in lobe after performing by k-means



Figure 13 The vessel image in lobe after performing connected components

Through by GMM processing, we get two clusters data parameters. These two parameters are the mean value and variance of each clustering, respectively. Because the images we used are injected contrast medium, we only keep the cluster which has the highest mean value in the result and segment vessel parts with other object in the background. Due to the bone intensity in the image shows the similar as normal vessel intensity, the bone area will also be kept after performing GMM. For keeping main vessel parts, the bone area and those small vessels around main vessel parts will treat as the noise which has to be removed. Hence, we combine connected components method again to remove those small areas in the image. Figure 15 shows the resulted image after performing connected components.



Figure 14 The main vessel image



Figure 15 The image after performing

Subsequently, the image after performing GMM contains aorta (there are ascending aorta and descending aorta) with contrast medium. It will produce false contour in the aorta image when doctor captured the image. It will affect the precision rate for detecting pulmonary embolism. Therefore, we have to remove the false contours which are not belonging to pulmonary artery before performing vessel tracking. Figure 15 shows some objects, from up to bottom is ascending aorta, pulmonary artery and descending aorta. There are obvious gradient changes among all of them. Hence, according to the characteristic, we use image projection method to find the boundary value to achieve removing the aorta and keep the pulmonary artery goal. Figure 16 shows the flowchart of removing descending aorta. In the removing descending aorta part, after vertical projection, we obtain a histogram as Figure 17(a).

For raising the precision rate of the boundary, we use





the 1-D smoothing filter to smooth the histogram for removing some noise, showing as Figure 17(b). Subsequently, we calculate the gradient of the histogram and smooth it in order to set the better boundary to remove the descending aorta. The result shows on Figure 17(c) and Figure 17(d). Figure 18 shows the image after removing descending aorta and backbone.

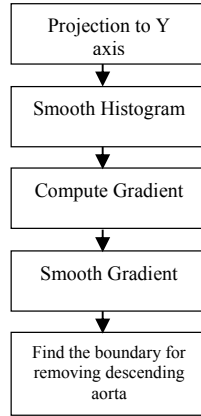


Fig. 16 The flow chart of removing descending aorta

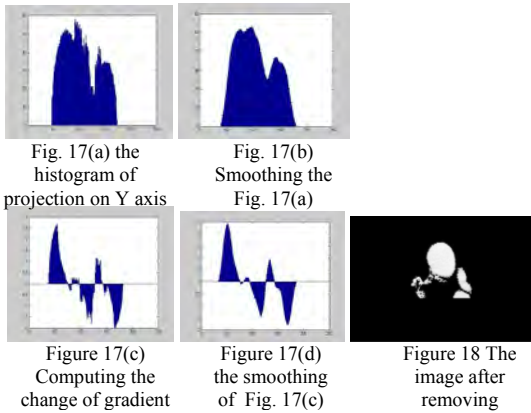


Figure 19 shows the algorithm of finding the boundary. The boundary value can separate the aorta and pulmonary artery. The input parameter array is the gradient array after smoothing.

```

Function boundary = count_boundary_dsc(array)
for j = R to 2
    value = array[j] * array[j-1];
    if value < 0
        boundary = j;
        break;
    end
end
end
  
```

Figure 19 The algorithm for finding the boundary

R is the position of starting searching. The position of R will change according to the different target. Assume W is the length of the gradient array, it is also the width of Y axis of the image. Due to the goal in here is removing the descending aorta, the calculation of starting searching position is  $R = (7 \times W) / 10$ . At the big change of gradient, the forward value and the backward are in different quadrants. Therefore, the values in different quadrants are multiplied to get the result which is lower than zero

as the characteristic. The characteristic can find the boundary value. Then, we use the image shown on Figure 18 to perform removing ascending aorta. Fig. 22 shows the flow chart of removing ascending aorta. Our method makes the own mask of each image. The mask can remove the ascending aorta in the image. In the removing ascending aorta part, we need to perform twice boundary computation, one is on Y axis and the other is on X axis. Figure 20 and Figure 21 show the algorithm of removing ascending aorta. First, we use Equation 3 to perform projection on Y axis and use Equation 4 to perform smoothing on histogram. Then we use Equation 5 to find the boundary. The result shows on Figure 23. Subsequently, we use Figure 23 to perform projection on X axis, and then perform as above steps to smooth the histogram and calculate the boundary. The result shows on Figure 24.

$$Vproj(y) = \sum_{x=1}^m Ip(x, y)$$

$$\text{where } y = 1, 2, \dots, n \quad (3)$$

$$Ip(x, y) = 0 \quad \text{if } I(x, y) = 0$$

$$Ip(x, y) = 1 \quad \text{if } I(x, y) \neq 0$$

$$Smooth\_Vproj(y) = \frac{1}{p+1} \sum_{i=y-\frac{p}{2}}^{y+\frac{p}{2}} Vproj(i), \quad (4)$$

$$\text{where } y = (p/2) + 1 \dots n - (p/2)$$

$$G(y) = (smooth\_Vproj[y+1] - smooth\_Vproj[y-1]) / 2,$$

$$\text{where } y = 2, \dots, n-1$$

$$G(y) = smooth\_Vproj[y+1] - smooth\_Vproj[y], \quad (5)$$

$$\text{where } y = 1$$

$$G(y) = smooth\_Vproj[y] - smooth\_Vproj[y-1]$$

$$\text{where } y = n$$

```

Function boundary = count_boundary_asc1(array)
for j = R to W-1
    value = array[j] * array[j+1];
    if value < 0
        boundary = j;
        break;
    end
end
end
  
```

Figure 20 The algorithm for finding the boundary on Y axis

The definition of  $R = (3 \times W) / 8$ , which represents the starting searching position, is the same as  $C = H / 2$  in the Fig. 20 and Fig. 21. And it will not be fixed according to the different target. The w is the length of gradient array, it is also the width on Y axis in the image. Here H is the length of the array, it is also the width on X axis in the image. Finally, we use the result of Fig. 24 as a mask to perform XOR operation with the corresponding position of original image. That will achieve the goal of removing ascending aorta. The result shows on Figure 25.

## 2.4 Vessel Tracking

In vessel tracking, we use the vessel distribution chart of pathology as the vessel skeleton chart. The vessel skeleton chart uses to track the alignment of the main



vessel.

```

Function boundary = count_boundary_asc2(array)
for j = C to H-1
    value = array[j] * array[j+1];
    if value < 0
        boundary = j;
        break;
    end
end

```

Figure 21 The algorithm for finding the boundary on X axis

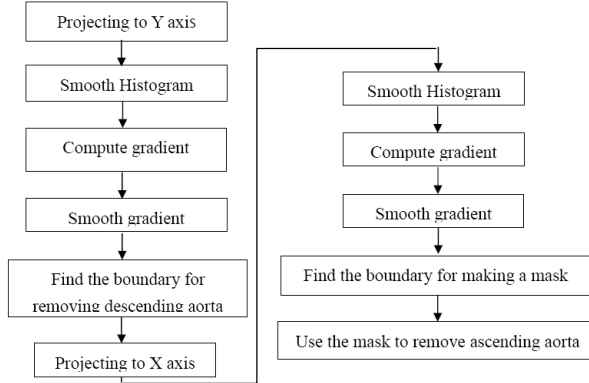


Figure 22: The flow chart for removing ascending aorta



Fig. 23 The first result of performing finding boundary



Fig. 24 The second result of performing finding boundary



Fig. 25 The image after removing ascending aorta

Due to the thickness of slices we used is 1mm, we can use such the details from slices to observe the change and build the vessel skeleton chart. After a series of background removing and backbone and aorta removing, there is pulmonary artery in the image. We know that the pulmonary embolism exists in pulmonary artery, so we perform pulmonary artery tracking and pulmonary embolism detection on the current image. In order to track becoming easier, we perform a preprocessing step before tracking pulmonary artery. Fig. 26 shows the flow chart of preprocessing for tracking pulmonary artery. According to the doctor explanation, the pulmonary artery will grow into “人” shape in 2D CTA image. In this part, we perform projecting on X axis and find the bigger change in gradient of the image. For letting tracking pulmonary easily and faster, we divide the pulmonary artery into two parts. Figure 27 shows the result image after projecting and dividing. Figure 27(a) shows the image which is removed background noise, backbone and aorta removing. Figure 27(b) shows the left side of pulmonary artery and Figure 27(c) shows the right side of pulmonary artery.

Then, we use the thinning method of image processing to find the skeleton of the image after dividing. Here we will introduce the thinning method. Let a object as  $O$ , and its boundary is  $B$ . If there is a pixel  $t$ ,  $e_1$  and  $e_2$  which  $t$  is in the object  $O$  and  $e_1$ ,  $e_2$  belong to boundary  $B$ .

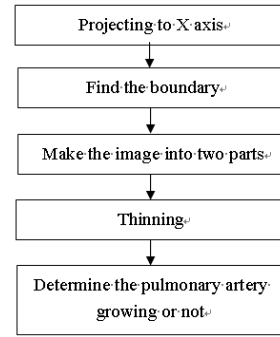


Figure 26 The preprocessing flow chart of pulmonary artery tracking



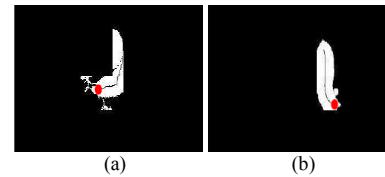
(a) (b) (c)

Figure 27 The pulmonary artery image. (a) Before performing projection and segmentation; (b) After performing projection and segmentation of the left pulmonary artery image; (c) After performing projection and segmentation of the right pulmonary artery image

If all of them can satisfy the distance function  $d$  as eq. 6,  $t$  can be regarded as the element of the skeleton of the object  $O$ .

$$d(t, e_1) = d(t, e_2) \quad (6)$$

Here  $d(t, e_i)$ ,  $1 \leq i \leq 2$  represents the distance between  $t$  and  $e_i$ . In the part of distance function, we can get the different skeleton according to the different distance function. Here we use is Euclidean distance function to get the skeleton of pulmonary artery. The result shows on Figure 28. The black line represents the skeleton.



(a) (b)

Figure 28 The image contains pulmonary artery and its skeleton. (a) The skeleton of left pulmonary artery ; (b) The skeleton of right pulmonary artery

Finally, we use the skeleton of pulmonary artery to track the pulmonary artery in the image whether the image has normal growing or not. Figure 29 shows the flow chart of pulmonary artery tracking. When getting the skeleton of pulmonary artery; meanwhile, we can obtain the endpoints of each skeleton. In Figure 28, the red point is the endpoint of skeleton. According to the image we used is a sequence 2D image, the image in time  $T$  and time  $T+1$  will not have big changed. Hence, we use endpoints as the foundation to perform pulmonary artery tracking. Before tracking pulmonary artery, we will check a record file which records the position of endpoints.





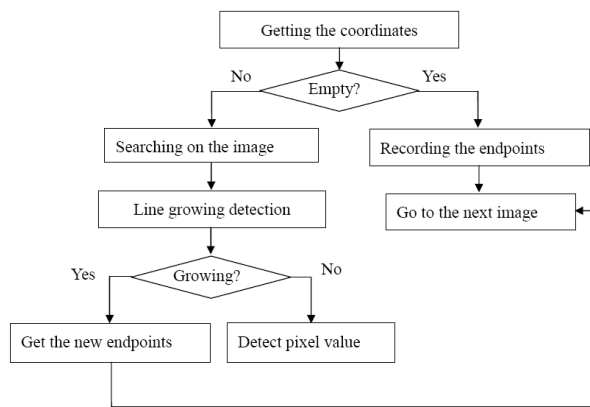


Figure 29 The flow chart for tracking pulmonary artery

If there does not have the endpoint records of forward image, we will record the endpoint of the current image to terminate tracking pulmonary artery and proceed processing for the next time image continually. On the contrary, if there is the endpoint information, then use this position to collate with the same position of the current image and detect the skeleton in the region whether it grows or not. If the skeleton grows, the system records the newest endpoint and detects whether it has the embolism characteristic in the new growing path. Otherwise, the system also detects the surrounding area of endpoint to understand whether it exist the feature of embolism or not.

Finally, according to the professional doctor's explanation, the CT value of embolism will not be variation change with different person. There is a characteristic about the CT value of embolism; that is, it has a fixed range. From the experiment, the CT value range of embolism is from 1050 ~ 1100 HU. We use the CT number range of embolism to be feature for detecting embolism in main and branch vessel.

### 2.5 Evaluation

Here, we use the template matching method to achieve precision computing. The professional doctor will select the embolism places in the image to produce the grown truth image for all of the images. These grown truth images will be regarded as template images. The produced image of our system will be regarded as comparison image. The comparison image will compare with the ground truth image. And then, according to the comparison result, the precision rate will be calculated. The matching we used adopts blob-by-blob method to proceed. And then, we use the result of the output to take as the system performance.

## 3. Conclusion

### A. The introduction of data set

The instrument to get the image is GE Fast Helical CT. The data set we used all contain pulmonary embolism. The image is represented by 12 bits. The range of CT value of the image is from -2048 to 2048. We adjust the range of CT value into all of the positive values. Hence, the range will become from zero to 4095. The CTA image data is 2D CTA image data. The slices of the image is 1mm, the resolution of the image is  $512 \times 512$ . Each data

set contains 250 ~ 280 images, deducting those images without pulmonary artery; the real number of images is about 30 ~ 50

### B. Experimental results

Herein, we will divide the result into two parts to explain. First, we use twenty questionnaires to obtain doctor's operation suggestion. Next, through by the professional doctor's knowledge and suggestion, we adjust the system parameter. For example, the parameter of disk size of closing operator sets 2 and the execution time sets 1, we can get the better result. Figure 30 shows the CTA image with pulmonary embolism. Additionally, the branch vessels in lobe are hard to observe. The doctor also inconveniences to detect embolism on those branch vessels. But, our method can detect the pulmonary embolism in the branch vessel. Figure 31 shows the result.

Here, we enumerate some experimental results, the left images in the Figure 32 and Figure 33 are grown truth images and the right images are our experimental images. Fig. 32 shows the successful cases and Fig. 33 shows the failure cases. The red circles in the grown truth images are embolisms which are selected by the professional doctor. The white rectangles in the system images are embolisms which are selected by our system.

In the Figure 32, our system detects the embolisms position correctly. But in the Figure 33, some human tissues may cause this kind of error. For example, there are two embolisms in the pulmonary artery in the upper images of Figure 33. Because the first embolism does not stick the pulmonary artery entirely, our system loses the information of the first embolism; the bottom image shows the soft tissue in the intermediate pulmonary artery. As the experiment results, there are similar CT value of soft tissue, fat and embolism, our system will produce the error on these kinds of images.

Subsequently, we use TPR (or sensitivity), FNR and Precision (p) index to present recognition result. When the TPR and p value are higher and FNR value is lower, the system is better. Table 2 shows the performance of our system with using the CTA images from 16 patients to detect the embolism in the pulmonary artery. The average precision rate is about 0.83. From the above table, our system shows the lower TPR value but higher FNR value. That is because this experiment only uses pulmonary artery tracking to detect the embolism. If the pulmonary artery is in the growing stage and grows to the embolism not yet, our system cannot perform embolism detection. Table 3 shows the performance of our system in the branch vessel detection. The average precision rate is about 0.826.

Finally, we found some doctors to use our system. We use questionnaire method to get doctor's opinions. From the result of questionnaire analysis, we discover that there are 90% doctors to be willing used our system.

In this paper, we propose an automatic pulmonary embolism detection system. We will divide the problem into two parts to discuss. One is the main vessel processing. The other is branch vessel processing. They have different processing method. From experiment result, we discover that our proposed method has 83% successful rate and can find the pulmonary embolism in



branch vessel. The successful rate in branch vessel detection is 82.6%. The system will efficient relieve doctor's loading.

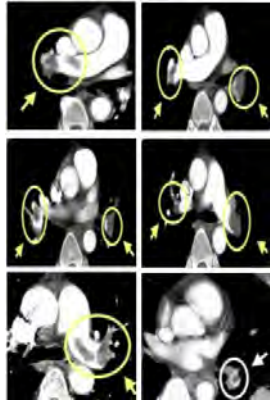


Figure 30 The images with pulmonary embolism.

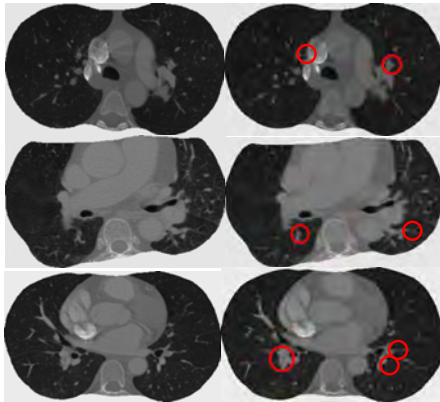


Figure 31 The left images are original image and the right images are the image with pulmonary embolism in branch vessel.

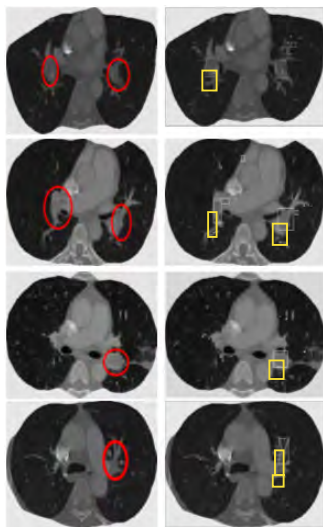


Figure 32 The experiment results (Successful cases)

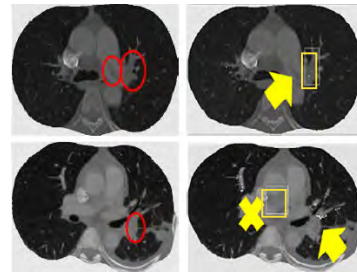


Figure 33 The experimental results (Fallacious cases)

Table 2: The performance of our system in main vessel

Data set No.	<i>TPR</i>	<i>FNR</i>	<i>p</i>
A patient	0.5	0.5	0.75
B patient	0.58	0.42	0.82
C patient	0.53	0.47	0.79
D patient	0.6	0.4	0.84
E patient	0.57	0.43	0.82
F patient	0.47	0.53	0.78
G patient	0.55	0.45	0.83
H patient	0.66	0.34	0.85
I patient	0.7	0.3	0.9
J patient	0.68	0.32	0.87
K patient	0.63	0.37	0.86
L patient	0.56	0.43	0.82
M patient	0.52	0.48	0.8
N patient	0.71	0.29	0.91
O patient	0.5	0.5	0.75
P patient	0.6	0.4	0.84

Table 3: The performance of our system in branch vessel

Data set No.	<i>TPR</i>	<i>FNR</i>	<i>p</i>
A patient	0.55	0.45	0.80
B patient	0.60	0.40	0.84
C patient	0.47	0.53	0.75
D patient	0.62	0.38	0.86
E patient	0.40	0.60	0.72
F patient	0.48	0.52	0.76
G patient	0.56	0.44	0.83
H patient	0.66	0.34	0.85
I patient	0.70	0.30	0.9
J patient	0.68	0.32	0.87
K patient	0.63	0.37	0.86
L patient	0.58	0.42	0.83
M patient	0.52	0.48	0.81
N patient	0.75	0.25	0.95
O patient	0.51	0.49	0.77
P patient	0.61	0.39	0.81



#### 4. Acknowledgements

The authors would like to thank the department of medical imaging, Chung Shan Medical University Hospital, Taichung, Taiwan, for medical image providing and help.

#### 5. References

- [1] Bouma, H. Sonnemans, J.J. Vilanova, A. Gerritsen, "Automatic detection of Pulmonary embolism in CTA images," IEEE Transaction on medical imaging, Vol. 28, pp. 1223-1230, 2009.
- [2] Michiels JJ, Hoogsteden H and Pattynama PM, "Non-invasive diagnosis of pulmonary embolism," Acta Chir Belg, Vol. 1, pp. 26-34, 2005.
- [3] P. Stein, S.E. Fowler, L.R. Goodman e.a, "Multidetector computed tomography for acute pulmonary embolism," New Engl. J. Med., Vol. 354, pp. 2317-2327, 2006.
- [4] U. Joseph Schoepf and Nicolaus Holzkecht, "Subsegmental pulmonary emboli: improved detection with thin-collimation multi-detector row spiral CT," Radiology, pp. 483-490, 2002.
- [5] Reinhard Groell and Karl H. Peichel, "Computed tomography densitometry of the lung: a method to assess perfusion defects in acute pulmonary embolism," European Journal of Radiology, Vol. 32, pp. 192-196, 1999.
- [6] Joachim E. Wildberger and U. Joseph Schoepf, "Approaches to CT perfusion imaging in pulmonary embolism," Seminar Roentgenol, Vol. 40, pp. 64-73, 2005.
- [7] Peter Herzog and Joachim E. Wildberger, "CT perfusion imaging of the lung in pulmonary embolism," Academic Radiology, Vol. 10, pp. 1132-1146, 2003.
- [8] Wildberger JE, Niethammer MU and Klotz E, "Multi-slice CT for visualization of pulmonary embolism using perfusion weighted color maps," Fortschr Rontgenstr, Vol. 173, pp. 289-294, 2001.
- [9] Simon J. Scurrrell, "Automatic detection of pulmonary embolism using computation intelligence," A dissertation of university of the Witwatersrand, October 2006.
- [10] Schoepf UJ and Costello P, "CT angiography for diagnosis of pulmonary embolism: state of the art," Radiology, Vol. 230, pp. 329-337, 2004.
- [11] A. Frigyesi, "An automated method for the detection of pulmonary embolism in V/Q scans," Medical Image Analysis, Vol. 7, No. 7, pp. 341-349, 2003.
- [12] R. Damper and I. Middleton, "Segmentation of magnetic resonance images using a combination of neural networks and active contour models," Medical Engineering & Physics, Vol. 26, No. 1, pp. 71-86, January 2003.
- [13] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing," Addison-Wesley Publishing Company, 1993.
- [14] Margarida Silveira and Jorge Marques, "Automatic segmentation of the lungs using multiple active contours and outlier model," IEEE Engineering in Medicine and Biology Society Annual International Conference, pp. 3122-3125, 2006.
- [15] D. Xu, J. Lee, D.S. Raicu, J.D. Furst and D.Channin, "Texture Classification of Normal Tissues in Computed Tomography," The 2005 Annual Meeting of the Society for Computer Applications in Radiology, Orlando, Florida, June 2-5, 2005.
- [16] Park W., Hoffman E.A. and Sonka M, "Segmentation of intrathoracic airway trees: a fuzzy logic approach," IEEE Transaction on Medical Imaging, vol. 17, No. 8, pp. 489-197, Aug. 1998.
- [17] Blechschmidt R.A., Werthschutzky R and Lorcher U, "Automated CT image evaluation of the lung: a morphology-based concept," IEEE Transactions on Medical Imaging, Vol. 20, No.5, pp. 434-442, 2001.
- [18] Dajnowiex, M. and Alirezaie, J., "Computer simulation for segmentation of lung nodules in CT images," IEEE International Conference on Systems Man and Cybernetics, Vol. 5, pp. 4491-4496, 2004.
- [19] Yongbum Lee, Takeshi Hara, Hiroshi Fujita, Shigeki Itoh and Takeo Ishigaki, "Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique," IEEE Transactions on Medical Imaging, Vol. 20, No. 7, pp. 595-604, 2001.
- [20] Augusto Silva, Jose S. Silva, Beatriz S. Santos and Carlos Ferreira, "Fast Pulmonary Contour Extraction in X-ray CT images: A Methodology and Quality Assessment," Medical Imaging 2001: Physiology and Function from Multidimensional Images, Vol. 4321, pp. 216-224, 2001.
- [21] Hu s., Hoffman E. A., Reinhardt and J. M., "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," IEEE Transactions on Medical Imaging, Vol. 20, No. 6, pp. 490-498, 2001.
- [22] Shojaii, R. Alirezaie and J. Babyn, P., "Automatic lung segmentation in CT images using Watershed Transform," IEEE International Conference on Image Processing, Vol. 2, pp. 1270-1273, 2005.
- [23] U.J. Schoepf, N. Holzkecht, T.K. Helmberger, et al., "Subsegmental pulmonary emboli: improved detection with thin-collimation multi-detector row spiral CT," Radiology, Vol.2, pp. 483-490, 2002.
- [24] Y. Masutani, H. MacMahon and K. Doi, "Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis," IEEE Transactions on Medical Imaging, Vol. 21, Issue 12, pp. 1517-1523, 2002.
- [25] C. Zhou, et al., "Computerized detection of pulmonary embolism in 3D computed tomographic (CT) images: vessel tracking and segmentation techniques," In Medical Imaging 2003: Image Processing. Proceedings of the SPIE, Vol. 5032, pp. 1613-1620, May 2003.
- [26] E.A. Chiang, et al., "Detection of pulmonary embolism: comparion of paddlewheel and coronal CT reformations-initial experience," Radiology 228 (2), pp. 577-582, 2003.
- [27] A.P. Kiraly, D.P. Naidich, and C.L. Novak, "Cartwheel projections of segmented pulmonary vasculature for the detection of pulmonary



embolism," Medical Imaging 2005:Image guided procedures, and display. Proceedings of SPIE - Int'l. Society for Optical Engineering, Vol. 5733, Part 1, pp. 69-78, 2005.

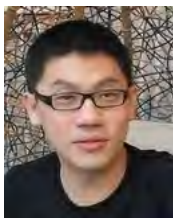
- [28] E. Pichon, C.L. Novak, A.P. Kiraly and D.P. Naidich, "A novel method for pulmonary emboli visualization from high-resolution CT images," SPIE Medical Imaging, Vol. 5367, pp.161-170, 2004.
- [29] A.P. Kiraly, C.L. Novak and D.P. Naidich, "A visualization-Based Method for Pulmonary Emboli Identification within High-Resolution CT Images," European Congress of Radiology, 2005.
- [30] A.P. Kiraly, "3D Image Analysis and Visualization of Tubular Structures," PhD thesis, The Pennsylvania State University, 2003.
- [31] A.P. Kiraly, D. P. Naidich and C. L. Novak, "2D display of a 3D tree for pulmonary embolism detection," Computer-aided Radiology and Surgery - CARS 2005 - Proceedings of the 19th Int'l. Congress and Exhibition, Vol. 1281, pp. 1132-1136, 2005.
- [32] Henri Bouma, Jeroen J. Sonnemans, Anna Vilanova and Frans A. Gerritsen, "Automatic Detection of Pulmonary Embolism in CTA Images," IEEE Transactions on Medical Imaging, Vol. 28, No. 8, August 2009.

## Biographies



Hao-Hung Tsai was born in Hsinchu, Taiwan in 1973. He received his medical doctor from the Faculty of Post-Baccalaureate Medicine Department, Kaohsiung Medical University, Kaohsiung, Taiwan, in 2003. He is a visiting staff of the department of medical imaging, Chung Shan Medical University

Hospital, Taichung, Taiwan from 2007 to now. His research interest is medical image of breast and chest.



Chiun-Li Chin was born in Taipei, Taiwan in 1975. He received his Ph.D degree in the electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2006, and his M.S. degree in electric engineering from Chung Hua University in 2000. He is an Assistant

Professor of the Department of Applied Information Sciences of the Chung Shan Medical University, Taichung, Taiwan. From Jan. 2008 to Jan. 2010, His research interests include adaptive signal processing, image and video recognition, medical image processing, and embedded system.



Kun-Ching Wang was born in Kaohsiung, Taiwan in 1976. He received his Ph.D degree in the electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2005, and his M.S. degree in electric engineering from Feng Chia University in 2000.

He is an Assistant Professor of the Department of Information Technology & Communication of the Shin Chien University, Kaohsiung, Taiwan. From Jan. 2007 to Jan. 2008, He was a Design Engineer of Digital Signal Processing Tech. Dept., Industrial Technology Research Institute, Chutung, Hsinchu, Taiwan. His research interests include adaptive signal processing, speech recognition, speech enhancement, and digital IC design.



Yung-Chih Cheng was born in Nantou, Taiwan in 1986. She received her Master degree in the Application of Information Sciences from Chung Shan Medical University, Taichung, Taiwan, in 2010. She is a Ph. D. student of the Department of Computer Science and Information Engineering of National Cheng Kung University,

Tainan, Taiwan. From July 2008, her research interest is medical image processing.







## A SEGMENTATION SCHEME BASED ON UNIFORM LBP AND MORPHOLOGICAL APPROACH

Anuradha.S.G<sup>1</sup>, K.Karibasappa<sup>2</sup>, B.Eswara Reddy<sup>3</sup>

<sup>1</sup>Department of Computer Science & Engineering, RYMEC, Bellary, Karnataka, India.

anuradha\_gagadin@rediffmail.com

<sup>2</sup>Department of Computer Science & Engineering, The Oxford College Bangalore, India.

k\_karibasappa@hotmail.com

<sup>3</sup>Department of Computer Science & Engineering, JNTUA, Anantapuramu, Andhra Pradesh, India

eswarcejntu@gmail.com

### Abstract

One of the essential tasks in a wide variety of applications and image processing domains is the need of automatic segmentation. The image segmentation divides an image into homogeneous regions mainly to locate objects and boundaries such as curves, lines, regions etc. The aim of image segmentation is to extract significant features of image, thus understanding, interpretation, description, analysis of the image scene becomes easy and useful for the machine to perform future tasks. Region segmentation divides images into regions based on pixel intensities, patterns, locations, local shapes and textures or combinations. The region segmentation mainly depends on how effectively the segmentation scheme captures the local attributes. This paper captures the local information by converting the grey level image into a uniform local binary pattern (ULBP) image. The advantage of this scheme is it only contains the significant information and all non significant local information is represented under a single label. Histogram equalization is used to enhance the local contrast. A morphological treatment is given to this, to smooth the regions and to close narrow gaps between two structures without growing the size of the structures. The Otsu's threshold is used for the final segmentation. The experimental results on four different databases demonstrate the success of the proposed method, compared to many other methods.

**Keywords:** Local attributes, Otsu threshold, region segmentation, histogram equalization..

### 1. Introduction

One of the most vital objectives of image analysis is image segmentation. This is because the succeeding steps of image analysis such as classification, illustration, description and image understanding and

restoration are mainly dependent on segmentation results. The image segmentation plays a critical role in a variety of pattern recognition applications such as robot vision, cartography, inspection of textile products, criminal investigation, remote sensing, object identification and recognition, military surveillance, quality assurance in industries, facial recognition and medical imaging etc. Image segmentation is performed usually using gray level intensities, color, texture, shape or any other feature of interest according to the particular application. The choice of tolerable segmentation idea is essential as it affects the segmentation system. Image segmentation [4, 6, 7, 8] divides an image into different regions depending on various attributes.

There are basically two types of image segmentation methods i) supervised ii) unsupervised. The unsupervised segmentation is more flexible for real world applications. And it is usually more computationally expensive. The problem of image segmentation has established significant concentration in the literature [10], [11], [12][19, 20, 22, 21]. There are various segmentation methods based on color of the image and it has wide applications in many areas [13] [14], [15] [16] [17] [20]. Various researchers proposed segmentation methods i.e. based on random field model [3], Pattern trends based on texture units [5], texton patterns to segment retinal vessels [2], three dimensional bi-directional histograms [1]. A segmentation method that is invariant in terms of shape, size, and intensity values is proposed for medical image applications [18]. The accuracy of segmentation is highly dependent on the success or failure of each computerized analysis procedure. A wide-range of research with various integrated methods have been reported in the literature for segmentation, however, still no method is found to be accepted and appropriate for all kinds of images. This





indicates clearly, all segmentation algorithms cannot be equally applicable to a certain application [9]. Recent studies have shown that outstanding texture classification and discrimination can be obtained with local binary patterns (LBP)[23,24,25,26,27]. To capture efficiently the local pattern of texture elements in this paper, we choose local binary pattern (LBP) operator as texture local feature descriptor.

The present paper organized as follows: The section 2 describes the related work. The section 3 and 4 describes the proposed method and results and discussions respectively. The section 5 describes the conclusion.

## 2. Related work

### 2.1 Mathematical morphology

Mathematical morphology (MM) is one of the popular non-linear theoretical models for image analysis and processing. Mathematical morphology is referred as a component of natural science and deals with various image components like shape, topology, connectivity etc. The morphological operations are derived from algebraic operators. Morphology is one of the oldest non linear theory and it is examined in 1960 by Matheron and Serra and it is an extension of Minkowski's set theory [28], [29]. Most of the non morphological image processing methods are mostly unsuccessful to find solution for the problems that deal with geometrical variations and aspects of the image. Morphological methods have advantages in dealing with textures due to their nonlinear nature. The morphological transformations retain basic topological properties of objects. Morphology plays a crucial role in segmentation where geometrical properties such as shape, size, connectivity or dissimilarity are considered as feature parameters. There are many morphological operators and many others are derived recently. The interesting thing of morphological operators is all of them are based on two simple operations i.e. dilation and erosion. Morphological operations can be implemented on binary and grey level image. They can also be defined with Euclidean (isotropic) or non-Euclidean (geodesic) metrics.

The morphological operators are very useful for boundary detection, image enhancement, image segmentation, image smoothing, preprocessing, and removal of noise, image understanding and analysis of images. They are widely used and preferred over linear approaches due to their simplicity, efficiency and mainly for direct geometric interpretation.

#### 2.1.1 Gray Value Morphological Processing

The present paper used morphological closing operation on grey level images and grey level morphological operations are described below. The techniques of morphological filtering can be extended to gray-level images. To simplify matters the present

study restricted the presentation to structuring elements,  $Q$ , that comprise a finite number of pixels and are convex and bounded. However, the structuring element has gray values associated with every coordinate position as does the image  $A$ .

The dilation,  $D_G()$  in grey level is given by equation 1.

$$D_G(P, Q) = \max_{[j,k] \in Q} \{P[m-j, n-K] - Q[j, k]\} \quad (1)$$

The erosion,  $E_G()$ , is given by the following equation 2.

Erosion

$$E_G(P, Q) = \min_{[j,k] \in Q} \{P[m+j, n+K] - Q[j, k]\} \quad (2)$$

For a given output coordinate  $[m,n]$ , the structuring element is summed with a shifted version of the image and the maximum/ minimum encountered over all shifts within the  $J \times K$  domain of  $Q$  is used as the result of dilation/erosion in grey level morphology.

The images grow and shrink in size by dilation and erosion operators respectively. The quantity, the type and the shape of the image grow or shrink depends upon the structuring element used in these operations. This obviously reflects the truth that morphological operations or processing is dependent on the size, shape and amount of structuring element used. Any morphological operation without a structuring element has no sense i.e. same as low pass filter of an image without specifying the filter.

The morphological grey level opening and closing are defined in equations 3 and 4.

$$\text{Opening} - O_G = D_G (E_G (P, Q) Q) \quad (3)$$

$$\text{Closing} - C_G = E_G (D_G (P, Q) Q) \quad (4)$$

The erosion followed by dilation is called morphological opening. In opening the erosion of an image removes all structures that the structuring element cannot fit inside. Further shrinks all other structures. Then by dilating the result of the erosion with the same structuring element, the structures that are survived by the erosion (were shrunken, not deleted) will be restored. The name tells that the operation can create an opening between two structures that are connected only in a thin bridge without shrinking the structures (as erosion would do). Opening generally smooths the contour of an object, breaks narrow isthmuses, and eliminates thin protrusions.

The dilation followed by erosion is called closing. The morphological dilation of an image expands the object and fills small gaps. By eroding this image will retain their structure and form, without any change, but tiny holes filled by dilation will disappear. Images merged by the dilation will not be separated again. The image will be smoothed by closing. This smoothing



usually mingles thin breaks and long thin gulfs. This removes or eliminates minute holes, and fills gaps in the contour. This retains the uniformity of a local region.

The advantage of grey level morphology is it reduces significantly the overall complexity of processing by the use of symmetric structuring elements. The definitions of above are reduced to the following equations 5, 6, 7 and 8.

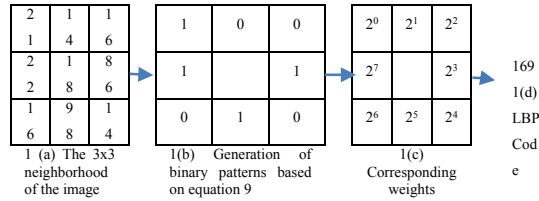


Figure 1: Local binary pattern approach.

$$Dilation - \\ D_G(P, Q) = \max_{[j,k] \in Q} \{P[m-j, n-k]\} = \max_Q(P) \quad (5)$$

$$Erosion - \\ E_G(P, Q) = \min_{[j,k] \in Q} \{P[m-j, n-k]\} = \min_Q(P) \quad (6)$$

$$Opening - O_G(P, Q) = \max_Q \left( \min_Q(P) \right) \quad (7)$$

$$Opening - O_G(P, Q) = \min_Q \left( \max_Q(P) \right) \quad (8)$$

The noteworthy conclusion from this the maximum and the minimum filter are exactly same as dilation and erosion in grey level respectively. In this case the shape of the filter window will become the specific structuring element. For a window of size J×K, the maximum or minimum filter is separable into two, one-dimensional windows. Further one can compute easily in incremental form, a one-dimensional maximum or minimum filter. This means the computational complexity per pixel of dilation or erosion is O(constant). Thus these morphological operations and also all other operations of morphology are independent of J and K.

## 2.2. Local binary pattern

The Local binary pattern (LBP) was introduced by Ojala et al. [30] in 1996. The LBP operator converts the grey levels of an image into integer codes called as LBP codes. The LBP code describes the local patterns, attributes and properties very significantly. The process of evaluating LBP codes are shown in Figure 1. Initially the neighborhood pixels are converted in to binary by computing the grey level differences between the central pixel and the corresponding neighborhood pixels based on as given in equation 9.

Then the binary values of the neighbors are multiplied with the corresponding weights. Summation of these gives a LBP code for the window and the central pixel value is replaced with the derived LBP code as given in equation 10.

$$b_i = S(P_c - P_i) \quad i=0 \text{ to } 7 \quad (9)$$

$$\text{where } b_i = \begin{cases} 1 & S(x) \geq 0 \\ 0 & S(x) < 0 \end{cases}$$

$$lc = \sum_{i=0}^7 b_i * 2^i \quad (10)$$

where  $P_c$ ,  $P_i$  represents the gray values of central pixel and neighborhood pixel  $i$ ,  $b_i$  represents the binary value of the corresponding neighboring pixel.  $lc$  represents the LBP code for the 3x3 window. The transformation process of LBP code is also shown in the Figure 1.

In literature uniform LBPs (ULBP) are derived on binary LBP. The ULBPs are considered as most important properties of image texture. More than 85 % of image windows will have ULBP's. The remaining windows are treated as Non-uniform LBP's (NULBP). The LBP codes are defined as uniform if they contain at most two circularly bitwise transitions from 0 to 1 or vice versa, and non-uniform patterns if otherwise [31]. While mapping a LBP image into ULBP image, unique code is used for each ULBP and all non-uniform patterns are accumulated with a single code and treating them as miscellaneous. Most LBPs in natural images, textures and human faces are uniform patterns [31, 32]. A 3 x 3 neighborhood with 8-neighbors derives 256 ( $2^8$ ) LBP codes, whereas by transforming this in to ULBP codes the number of codes will be reduced from  $2^8$  to  $P(P-1)+3$ . Here  $P$  represents the number of neighboring pixels. The present paper established, the advantage of considering ULBP's because it retains the basic fundamental properties of the image, which is very useful for segmentation purpose and it greatly reduces the image dimension.

## 2.3. Thresholding by Otsu method

One of the crucial steps in image processing is thresholding. Thresholding divides the image in to two or more regions based on the threshold levels. That's why thresholding is one of the crucial and significant steps in image segmentation. A threshold can be used to segment the image in the following way.

If  $G(x, y) > T$  then  $G(x, y) = 0$  else  $G(x, y) = 1$  or maximum intensity.

where  $G(x, y)$  represents the grey level intensity of the pixel at co-ordinate position  $x, y$ .  $T$  represents the threshold chosen randomly or by any other methods. To make segmentation more accurate and robust, the threshold should be selected automatically. Threshold is selected based on the familiarity and knowledge about the images and the application and also based on



intensity characteristics of the objects, sizes of the objects, fractions of an image occupied by the objects, number of different types of objects appearing in an image colour combination etc.

The Otsu method, as proposed by [33] is based on discriminate analysis. Otsu's method chooses the threshold by minimizing the within-class variance of the two groups of pixels separated by the thresholding operator. A measure of region homogeneity is variance (The regions with high homogeneity will have low variance). The advantage of Otsu threshold is it does not depend on modeling the probability density functions.

The Otsu threshold is based on a bimodal distribution (foreground pixels and background pixels) of gray-level values. The Otsu threshold operation performs the division of image pixels into two classes  $C_0$  and  $C_1$  (e.g., objects and background) at gray level  $t$ , i.e.,  $C_0 = \{0, 1, 2, t\}$  and  $C_1 = \{t+1, t+2, \dots, L-1\}$ .

Let  $\sigma^2_B$ ,  $\sigma^2_W$  and  $\sigma^2_T$  be the between-class variance, within-class variance and the total variance, respectively. An optimal threshold is determined in the Otsu method by minimizing one of the following criterion functions with respect to:

$$\lambda = \frac{\sigma^2_B}{\sigma^2_W}, \eta = \frac{\sigma^2_B}{\sigma^2_T}, \kappa = \frac{\sigma^2_T}{\sigma^2_W} \quad (11)$$

The optimal threshold 't' is defined as

$$t = \text{ArgMin } \eta \quad (12)$$

$$\sigma^2_T = \sum_{i=0}^{L-1} [1 - \mu_T]^2 P_i, \mu_T = \sum_{i=0}^{L-1} [i P_i] \quad (13)$$

$$\sigma^2_B = w_0 w_1 (\mu_0 - \mu_1)^2 \quad (14)$$

$$w_0 = \sum_{i=0}^t P_i, w_1 = 1 - w_0 \quad (15)$$

$$\mu_1 = \frac{\mu_T - \mu_0}{1 - \mu_0}, \mu_0 = \frac{\mu_T}{w_0}, \mu_t = \sum_{i=0}^t (i P_i) \quad (16)$$

$$n = \sum_{i=0}^{L-1} n_i \quad (17)$$

$$P_i = \frac{n_i}{n}$$

Where  $n_i$  is the number of pixels with gray-level 'i' and 'n' denotes spatial resolution of the given image. Moreover,  $P_i$  is the probability of occurrence of gray level 'i'.

The class probabilities  $w_0$  and  $w_1$  specify parts of the areas engaged by classes  $C_0$  and  $C_1$  for a chosen threshold 't'. The approximation of mean levels of the classes of original image is provided by class means  $\mu_0$  and  $\mu_1$ . The  $\eta^*$  represents the obtained maximum value of  $\eta$ . The  $\eta^*$  is used to compute the separability of classes  $C_0$  and  $C_1$  in the original or the bimodality of the histogram. This is crucial measure. It is uniquely determined within the range  $0 \leq \eta \leq 1$ . If and only if a given image shows a single constant gray level, then the lower bound (zero) is obtained. And the upper bound (unity) is achieved if and only if two-valued images are given. The proposed method used Otsu thresholding scheme to properly establish boundaries without noise and falling edge problems.

### 3. Methodology: Morphological Uniform LBP (MULBP) Segmentation Method

The proposed MULBP segmentation method divides the entire process into five steps. The first step converts the colour image into grey level image using RGB colour model. To represent effectively the local attributes of the image LBP is applied on the image. The image is encoded with Uniform LBP codes (0 to 59). Histogram equalization is applied to amplify the grey levels of ULBP coded image. This in fact enhances the local grey levels. Then morphological closing operation is applied to treat the uniform local regions. Finally Otsu threshold is applied to obtain segmented image. The proposed MULBP segmentation is a novel method because it achieves a good segmentation with five steps and thus it is more suitable to real time applications. The block diagram of proposed MULBP method is given in Figure 2.

STEP 1: Converts the colour image in to grey level image, by using RGB colour model.

STEP 2: To establish local features on a 3 x 3 neighborhood, the grey level image is converted in to LBP coded image. If the neighborhood forms a NULBP window then its central pixel is replaced with the miscellaneous code 59. The ULBP codes are quantized from 0 to 58. The formation of LBP codes and ULBP is given in section 2.2.

STEP 3: The histogram equalization (HE) is derived on the ULBP image of step 2, to identify interior details and to enhance the contrast of the image.

STEP 4: On the histogram equalized ULBP image, the morphological closing is applied to derive the uniform local regions of the image. Closing connects the objects that are close to each other. That is the closing fills up small gaps and smoothes the outline of the object. The morphological operations are described in section 2.1.

Step 5: To establish boundaries in the image, Otsu thresholding is performed on the image obtained in step 4 as explained in section 2.3.

The proposed MULBP evaluated the performance of the segmentation scheme by deriving various statistical measures on segmented images. One of the popular methods to evaluate the performance of the segmentation method is by subjective evaluation. In this a human visually approximates by comparing the segmentation results with various other segmentation approaches. The supervised evaluation is the other method for assessment. In this segmented output, will be compared with the original image. Both the above assessment methods need interaction of the user, that's why they are impractical in several computer vision tasks. Therefore to identify the performance of the segmentation process "unsupervised approaches" [34, 35] are required. The comparison with ground truth or original images is not required in unsupervised



estimation methods, thus it takes less time in evaluating the performance of the segmentation method. That's why these estimation methods are popular and preferred in real time and other applications.

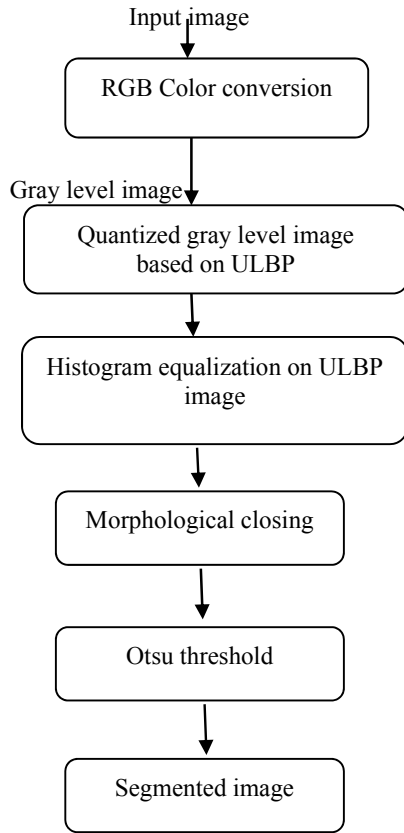


Figure 2: Block diagram of the proposed MULBP method.

The present paper considered the following popular and widely referred segmentation metrics i.e. standard deviation, contrast, discrepancy and entropy. Discrepancy is given as

$$Discrepancy = - \sum_i^{I_h} \sum_j^{I_w} (X_{gl}(i,j) - Y(i,j)) \quad (18)$$

Where  $X_{gl}(l, m)$  and  $Y(l, m)$  represents the gray level values at pixel co-ordinates  $l, m$  of the original and final image. If the value of discrepancy is high, then it indicates a better segmentation.

Entropy of an image is given as

$$Entropy = - \sum_i \sum_j S(i,j) \log(S(i,j)) \quad (19)$$

Over segmentation and under segmentation can be assessed with entropy values. An entropy value less than 1 and above 1.5 represents over segmentation and under segmentation respectively.

Standard deviation of a given vector is expressed as

$$S = \left[ \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{\frac{1}{2}} \quad (20)$$

Where  $x_i$  and  $\bar{x}$  are the value of vector and average of all values.

A better segmentation is estimated with lower values of standard deviation. Internal region contrast is defined as

$$I_j = \frac{1}{S_j} \sum_{a \in R_j} \max\{c(a, b), b \in W(a) \cap R_j\} \quad (21)$$

Where neighborhood is denoted by  $W(a)$ , and  $c(a, b) = |C_x(a) - C_x(b)|$  is the contrast of pixel 'a' and 'b'. The region uniformity is measured by internal contrast,  $I_j$ . The  $I_j$  represents "region average Max Contrast". Uniformity is one of the important attributes of the segmentation. A good segmentation method should divide the image without disturbing the region uniformity. A high uniformity is measured with a low value of internal contrast.

#### 4. Results and discussions

The proposed MULBP segmentation method is tested on four large databases namely WANG [39], OXFORD flowers [40], INDIAN facial expressions [41] and standard images from Google (Lena, Camera man, House, Mandrill, and Ship). We have tested the MULBP segmentation method on 150 images from each of the above data bases and this result to a total of  $4 \times 150 = 600$  different images.

The WANG database consists of 1,000 natural images. These images are manually selected from the famous Corel stock photo database. The WANG database divided these 1000 images in to 10 classes and each class contains 100 images. This database is selected because mostly these images are used for image retrieval and classification experiments. The good segmentation of these images will have further applications in retrieval and classification. The OXFORD flower database consists of flowers of 17 categories. The flower images are most suitable for segmentation experiments' because they contain sharp edges with different shapes and different local attributes. In each category there are 80 images. This results to a total of  $17 \times 80 = 1360$  images. The segmentation plays a crucial role in facial image recognition, image retrieval, facial expression identification, ageing problems and age classification. That's why the present paper chosen Indian data base that contains facial images. These images are taken at IIT Kanpur campus and placed in the web in February 2002. This set consists of 59 different classes and each of this contains 11 different facial images ( $59 \times 11 = 649$  Images). Some additional photographs are included in some classes. On a bright homogeneous background these images are captured. The resolution of images is  $640 \times 480$  pixels, with 256 grey levels per pixel. The facial images contains both genders male and female), with different age groups, with different expressions (neutral, smile, laughter, sad/disgust) and emotions.

To show the step wise performance, the proposed MULBP is applied on eight input images from the above data bases and results are shown in Figure 3, and they clearly establish the following facts.





1. The ULBP captured local information significantly.
2. The contrast of the local ULBP regions is enhanced by histogram equalization.
3. The proposed morphological treatment filled the small holes and established uniform local regions for a better segmentation.
4. The Otsu threshold established local boundaries of the image effectively and also removed unwanted non significant portions of the image.

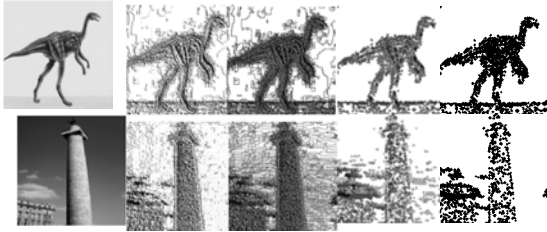


Figure 3.1: Images from WANG database.

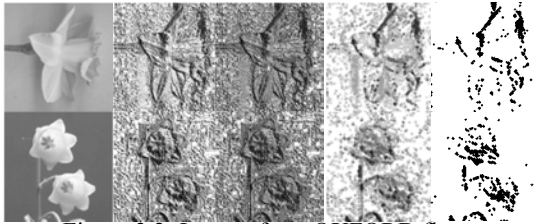


Figure 3.2: Images from OXFORD flower database.



Figure 3.3: Images from INDIAN database.

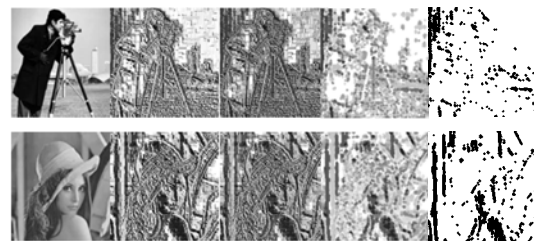


Figure 3.4: Images from Google database.

(a) (b) (c) (d) (e)

Figure 3: (a) the gray scale images (b) ULBP images of (a) (c) Histogram equalization outputs of (b) (d) Output images after closing operation (e) the MULBP segmented images.

The present paper evaluated the discrepancy, standard deviation, entropy and contrast values on the 600 segmented images (150 from each database) on the

proposed MULBP and other existing segmentation methods i.e. ISLGHEM [36], automatic threshold method [37] and wavelet based watershed method [38]. The average value of 150 images from each database is plotted in the form of bar graphs in Figure 4, Figure 5, Figure 6 and Figure 7.

The graphs clearly indicate the higher discrepancy value and a lower value of the standard deviation for the MULBP method over the other methods. This reflects a better segmentation of the proposed over the existing methods. The low value of internal contrast of MULBP represents a high uniformity in the region. A segmentation mechanism may lead into over and under segmentation. An under segmentation where as too many regions leads to over segmentation. A over segmentation is a undesirable characteristic. The over segmentation problem is overcome by the proposed method and it is reflected in the form of entropy value.

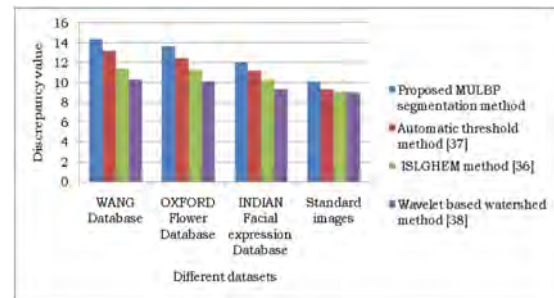


Figure 4: Discrepancy graph of various segmentation schemas on considered databases.

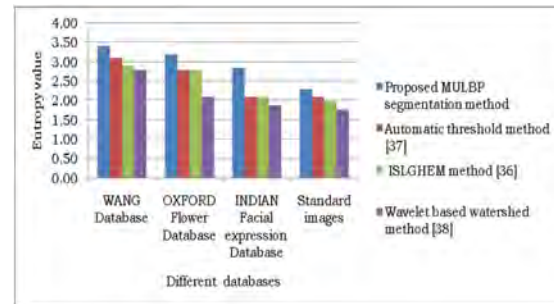


Figure 5: The entropy graph on segmentation methods using different databases.

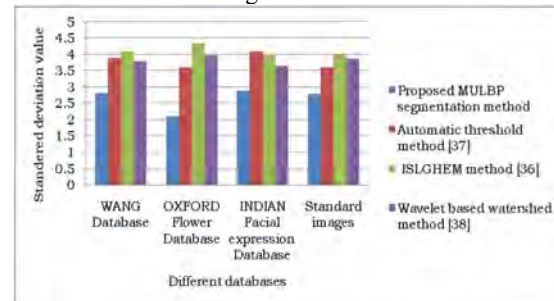


Figure 6: The standard deviation graph on segmentation methods using different databases.





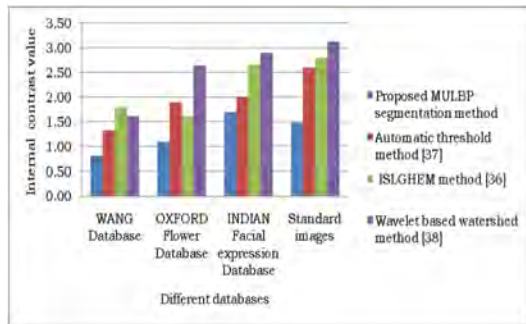


Figure 7: The internal region contrast graph on segmentation methods using different databases.

## 5. Conclusions

The MULBP identified local attributes efficiently by the usage of the powerful, simple and robust local operator i.e LBP. The other significant advantage of considering ULBP is it is rotationally invariant because the ULBP is measured on circular bit pattern of a 3 x 3 neighbourhood. The morphological treatment filled the small holes and connected borders of regions for a better segmentation. The Otsu threshold established local boundaries efficiently and provided better contrast and also removed the unwanted local scenes of the image. The whole process of segmentation is autonomous and requires no supervision. The MULBP segmentation improves the contrast of sharp details in light and dark areas because of LBP-window based property and it showed higher performance when compared other three methods of segmentation. The present method is simple and suitable to real time applications because it achieved good segmentation with five basic steps.

## References

- [1] Cula.O.G and Dana.K.J, "3D texture recognition using bidirectional feature histograms", in *Int. Journal Comp. Vis.*, vol.59, pp.33–60, 2004.
- [2] Donald.A, Adjero and Umasankar Kandaswamy, "Texton-based segmentation of retinal vessels", *Journal of Optical Society of America*, vol. 24, no. 5, pp.1384–1393, May 2007.
- [3] Idrissi sidi yassine, Samir belfkih, "Texture image segmentation using a new descriptor and mathematical morphology", in *Int. Arab Journal of Information Technology*, Vol.10, No.2, pp. 204-208, March 2013.
- [4] Kekre.H.B, Saylee Gharge, "Texture Based Segmentation using Statistical Properties for Mammographic Images", *Int. Journal of Advanced Computer Science and Applications(IJACSA)*, Vol. 1, No. 5, pp. 102-107, Nov 2010.
- [5] Kezia.S, ShantiPrabha.I, VijayaKumar.V, "Innovative segmentation approach based on LRTM", *Int. Journal of Soft Computing and Engg.*, vol. 2, no. 5, pp. 229-233, Nov. 2012.
- [6] Lutz Goldmann, Tomasz Adamek, Peter Vajda, "Towards Fully Automatic Image Segmentation Evaluation", in *Advanced Concepts for Intelligent Vision Systems Lecture Notes in Comp. Science*, vol. 5259, pp. 566–577, 2008.
- [7] Singh.S and Sharma.M, "Texture experiments with Meastex and Vistex benchmarks," in *Proc.Inter. Conf. on Advances in Pattern recog.*, *Lecture Notes in Computer Science*, 2013.
- [8] Wang.B and Zhang.L, "Supervised texture segmentation using wavelet transform," in *Proc. of Int. Conf. on Neural Networks and Signal Processing*, vol. 2, pp. 1078–1082, 2003.
- [9] Ahmed R. Khalifa et al., "Evaluating The Effectiveness Of Region Growing And Edge Detection Segmentation Algorithms", *Journal of American Science*, 6(10), pp.580-587, 2010.
- [10] M. Wirth, D. Nikitenko and J. Lyon, "Segmentation of the Breast Region in Mammograms using a Rule-Based Fuzzy Reasoning Algorithm", *GVIP Journal Special Issue on Mammograms*, pp. 13-21, 2007.
- [11] N. R. Pal and S. Pal, "A review on image segmentation techniques" *pattern recognit.*, vol. 26, pp. 1277-1294, 1993.
- [12] P. K. Sahoo, S. Soltani, and A. K. C. Wong "SURVEY: A survey of thresholding techniques" *Comput. Vis. Graph. Image process.* vol 41, pp. 233-260, 1988.
- [13] R. E. Cummings, P. Pouliquen, and M. A. Lewis, "A Vision Chip for Colour Segmentation and Pattern Matching", *EURASIP Journal on Applied Signal Processing*, no. 7, pp. 703-712, 2003.
- [14] Y. Yang, C. Zhen and P. Lin, "Fuzzy C-means clustering algorithm with a novel penalty term for image segmentation", *Opto-Electronics Rev.*, vol. 13, no. 4, 2005.
- [15] M. B. Meevathi and K. Rajesh, "Volterra Filter for color image segmentation", *Inter. Journal of Computer Science and Engineering*, vol. 2, no. 1, 2008.
- [16] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity", *IEEE Transaction on Image Processing*, vol. 9, no. 12, pp. 2071- 2082, 2000.
- [17] H. Seddik and E. Ben Braiek "Color Medical Images Watermarking, Based Neural Network Segmentation "GVIP Journal Special Special Issue on (Medical Image Processing), pp. 81-86, 2006.
- [18] An automated System for Liver CT Enhancement and Segmentation; M. Usman Akram, Aasia Khanum and Khalid Iqbal;



- ICGST-GVIP Journal, Volume 10, Issue IV, October 2010 .
- [19] Border Detection of Noisy Skin Lesions by Improved Iterative Segmentation Algorithm using LOG Edge Detector ;J.H.Jaseema Yasmin1, D. Muhammad Noorul Mubarak2 , M.Mohamed Sathik 3;;; ICGST-GVIP Journal, Volume 12, Issue 2, August 2012 .
- [20] Color Image Segmentation based on the Optimal Multilevel Thresholding Technique R. Harrabi and E. Ben Braiek;;; ICGST-GVIP Journal, Volume 12, Issue1, April 2012.
- [21] Efficient DIS Based Region Growing Segmentation Technique for VHR Satellite Images ;; Ashwini Kunte, Anjali Bhalchandra;; ICGST-GVIP Journal, Volume 10, Issue 3, August 2010 .
- [22] Objective Assessment of Nonlinear Segmentation Approaches to Gray Level underwater Images Zhengmao Ye;; ICGST-GVIP Journal, ISSN 1687-398X, Volume (9), Issue (II), April 2009.
- [23] V Vijaya Kumar, U S N Raju, K Chandra Sekaran, V V Krishna ,Employing long linear patterns for texture classification relying on wavelets, ICGST-Graphics, vision and image processing (ICGST-GVIP), Vol.8, No.5, pp. 13-21, Jan-2009.
- [24] V.Vijaya Kumar, U.S.N. Raju, K. Chandra Sekharan, V.V. Krishna , A new method of texture classification using various wavelet transforms based on primitive patterns, ICGST-Graphics, vision and image processing (ICGST-GVIP), Vol.8, No. 2, pp.21-27, July-2008.
- [25] Srikrishna. A. and Vijaya Kumar. V., Classification and recognition of digits by using mathematical morphology, SADHANA - Indian academy of sciences, Bangalore, Vol.35, No. 4, pp. 419-426, Aug-2010.
- [26] V.Vijaya Kumar, A. Nagaraja Rao, U.S.N. Raju, B.Eswar Reddy ,Pipeline implementation of new segmentation based on cognate neighborhood approach, International journal of computer science, Science publications, Vol.35, No.1, pp. 1-6, Feb-2008.
- [27] V.Vijaya Kumar, B. Eswar Reddy, A. Nagaraja Rao, U.S.N. Raju ,Texture segmentation methods based on combinatorial of morphological and statistical operations, Journal of multimedia (JMM), Academy publishers, Vol.3, No.1, pp.36-40, JMM, 2008,
- [28] J. Serra, Image Analysis and Mathematical Morphology. London, U.K.: Academic, 1982.
- [29] Introduction to mathematical morphology, Comput. Vis., Graph., Image Process., vol. 35, pp. 283–305, 1986.
- [30] T.Ojala, M. Pietikäinen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions,” Pattern Recognition, vol. 29, no. 1, pp. 51-59, 1996.
- [31] M. Inen, M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, Computer Vision Using Local Binary Patterns, vol. 40. New York, NY, USA: Springer-Verlag, 2011.
- [32] T. Ahonen, A. Hadid, and M. Pietikainen, “Face description with local binary patterns: Application to face recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
- [33] Otsu.N, "A threshold selection method from gray-level histograms", IEEE Trans. Sys., Man., Cyber, Vol.9, pp.62–66, 1979.
- [34] Cui.Y, Dong.H, Zhou.E.Z, “An Early Fire Detection Method Based on Smoke Texture Analysis and Discrimination”, Journal Congress on Image and Sig. Proc., pp.95–99, 2008.
- [35] Mounir Sayadi, Lotfi Tlig and Farhat Fnaiech, “A new texture segmentation method based on the fuzzy C-mean algorithm and statistical features”, Applied Mathematical Sciences, vol. 1, no. 60, pp. 2999 – 3007, 2007.
- [36] Saka.Kezia, Dr.I.Santi Prabha, Dr.V.VijayaKumar, A New Texture Segmentation Approach for Medical Images, International Journal of Scientific & Engineering Research, Volume 4, Issue 1, January-2013.
- [37] Yuan Been Chen and Oscal.T.C.Chen, "Image Segmentation Method Using Thresholds Automatically Determined from Picture Contents,” in Eurasip journal on image and video processing, vol.2009, pp. 1-16, 2009.
- [38] Yu-hua Chai, Li-qun Gao, and Shun Lu, Lei Tian, “Wavelet-based Watershed for Image Segmentation”, in Proc. of the 6th World Congress on Intelligent Control and Automation, Dalian, China, June 21 - 23, 2006.
- [39] Deselaers, T., Keysers, D., and Ney, H. 2008. Features for image retrieval: an experimental comparison. Inf. Retr. 11, 2 (Apr. 2008), 77-107. DOI=<http://dx.doi.org/10.1007/s10791-007-9039-3>.
- [40] Yuning CHAI, Victor LEMPITSKY, Andrew ZISSERMAN. BiCoS: A Bi-level Co-Segmentation Method for Image Classification. ICCV, 201.
- [41] Vidit Jain, Amitabha Mukherjee, 2002, The Indian Face Database, [http://vis-www. cs.umass. edu/~vidit/IndianFaceDatabase/](http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase/)



## Biographies



Anuradha.S.G graduated in B.E Computer Science and engineering in 2005 from Vijyanagara Engineering College, Bellary affiliated to Visvesvaraya Technological University (V. T. U.) Belgaum, Karnataka.

Completed M.Tech in the year 2007 in Computer Network Engineering from PG centre VTU, Belgaum and she joined as lecturer in department of CS&E RYMEC in 2007. Currently working as Assistant professor Department of CS&E, RYMEC, Bellary, and Karnataka. She is part time PhD research scholar at JNTUA in the CS&E department and carrying research work in the Image Processing area, published research articles at national and international level. She is a life member of ISTE professional society.



Dr.K.Karibasappa graduated in B.E, E&CE, Malnad College of Engineering, and Mysore University. He graduated his M.E in ETCE from Jadavapur University, Jadavapur Calcutta .He was

awarded PhD degree in Engineering from University college of Engineering sambalpur university, Burla, Orissa for the dissertation titled "Machine learning and Perception using Cognitive Methods" in the year 2004. His research interests include artificial intelligence, machine learning, image processing and pattern recognition .His has various paper published at national and international conferences and journal in the above areas .He is currently working as Principal The oxford college of Engineering, Bangalore, Karnataka, India.



Dr. B. Eswara Reddy Graduated in B.Tech.(CSE) from Sri Krishna Devaraya University in 1995. He received Masters Degree in M.Tech. (Software Engineering), from JNT University, Hyderabad, in 1999. He

received Ph.D in Computer Science & Engineering from JNT University, Hyderabad, in 2008. Currently, he is working as Professor of CSE Dept. & Vice Principal of JNTUA College of Engineering(Autonomous), Ananthapuramu. He has published more than 100 publications in national and international conferences and journals. His research interests include Pattern Recognition & Image Analysis, Data Mining, and Cloud Computing. He received UGC Major Research Project titled 'Cloud Computing Framework for Rural Health Care in Indian Scenario'. He is co-author of the text books 'Programming with Java' (Pearson/Sanguine) and 'Data Mining' (Elsevier India). He is a life member of CSI, ISTE, ISCA, Fellow-IE (India) and member of IEEE.







## Text lines Segmentation of Handwritten Arabic Script using Outer Isothetic Cover

Samia Snoussi, Fethi Ghazouani, Yosra Wahabi

Faculty of Computing and Information Technology Jeddah University, Saudi Arabia

RIADI Laboratory, ENSI, Manouba university, Tunisia

College of Community Service and Continuing Education, Umm Al Qura University, Saudi Arabia

Samia.maddouri@enit.rnu.tn, gfethi@yahoo.fr, yosra\_w@yahoo.fr

### Abstract

In this paper, we present a new bottom up segmentation method of handwritten Arabic texts into text lines. The proposed method uses the algorithm for construction of the Outer Isothetic Cover (OIC) of a digital object. This technique allows the construction of polygons of all connected components on the binary image of the document. The specificity of this method is that it allows the extraction of shape of the script and not a line as a geometric form. This will reduce overlapping between successive lines. The second specificity of this method is its independency from the variability of script size in the same document. The set of extracted polygons contains big and small ones. The big polygons correspond to Pieces of Words and the small ones correspond generally to diacritic points or punctuation mark and they are identified by their perimeters after a training step. Based on these results and added information like position, upper and lower limit of the components given by the OIC, we can extract lines from handwritten Arabic text. The proposed method is tested and evaluated on a sub-set of 100 randomly chosen handwritten Arabic texts selected from KHATT Database. The obtained result attempts 74% and needs to be improved. A new database composed by more than 1000 pages of printed and handwritten documents for tutorial and guiding of HAJJ steps is created (The Hajj is the pilgrimage to Mecca). We are working on the application of the proposed segmentation method on the HAJJ document database in order to extract desired information.

**Keywords:** *text Segmentation, handwriting, Outer Isothetic Cover, KHATT Database, HAJJ Database.*

### NOMENCLATURE

<b>OIC</b>	<i>Outer Isothetic Cover</i>
<b>KHATT</b>	<i>Arabic offline Handwritten Text database</i>
<b>RLSA</b>	<i>Run Length Smoothing Algorithm</i>
<b>OCR</b>	<i>optical character recognition</i>

### 1. Introduction

The segmentation is the most important phase in the applications of the pattern recognition. It allows preparing data for further processing steps like image processing, analysis and classification. In document processing field, the segmentation is essential for document recognition which it needs several steps of blocks classification, lines and words extraction from text blocks and normalization, words segmentation and features detection for each extracted information. This phase constitutes one of the major difficulties of all recognition system. Today, there is practically no reliable segmentation method for the handwritten Arabic document. The existing methods generally commit either "Over" or "Under" segmentation.

In this paper we propose a new method not to eliminate this segmentation problem, but we try to reduce it. We are interested in this work only on text document images. Essentially, for documents containing HAJJ and Umrah rules. The main application of this work is the segmentation of these documents. The identified segments will be transformed to text with OCR. Later, the obtained text will be analysed by a second intelligent system to extract HAJJ rules that will be used for automatic answers to questions about the HAJJ. This research was supported by the strategic Technologies Programs of the National Plan or Science, Technology and Innovation (MAARIFAH) in the Kingdom of Saudi Arabia. No: 13-INF134-05.

The segmentation of handwritten text is complicated by the variation of the interline distance and by the baselines undulation that often generates different orientations of the text. The characters in two adjacent lines may touch or overlap. This considerably complicates the text lines segmentation. In Arabic script, these situations frequently exist because of the presence of ascendant and/or descendant characters. On the other hand, the massive presence of diacritical symbols in Arabic script often generates false lines. In literature, most works of document segmentation are based on the decomposition of the image content into connected components. In this framework, we present





some works of text lines segmentation in order to justify the choice of our proposed method. Then, we present a new lines segmentation method of handwritten Arabic text. Afterward, we present the database used for the evaluation of our method. Finally, we show and discuss the obtained results.

## 2. State of the art

Several segmentation methods for handwritten documents have been carried in the literature [21], [22]. Some of these methods are interested in text line segmentation [23], [26], others in words or sub-words segmentation and certain in characters segmentation [25]. The segmentation methods can be roughly divided into three approaches: top-down, bottom-up and hybrid. Top-down methods consider the document in its globality and a single orientation is normally looked. Bottom-up methods are based on the analyses of elements of low level of the image like the pixels and the connected components [24].

### 2.1. Related Works

The projection profile technique is a top-down method most used for text line extraction [1][2]. Based on this technique, Arivazhagan et al. propose in [1], a static approach to line segmentation in handwritten documents, where the initial image is partitioned in vertical strip. Then, the histogram projection of every vertical strip is calculated. The first candidate lines are extracted among the first strips and the overlapped connected components are assigned into text lines by bivariate Gaussian densities. Also, Zahour et al. propose in [2] a partial projection-based method combined with slant detection and partial contour tracing to segment historical Arabic documents into text lines. The X-Y cut algorithm [3] based on projection profile is considered as top-down method. The principle of this method consists of cutting a binary image horizontally and vertically in several strips. Nicolaou and Gatos propose in [4] a top-down technique to segment handwritten documents image into text lines by shredding their surface with local minima tracers. After blurring the initial image in the goal to enhance text line areas, they segment the images surface along several white paths. In [5], Shi and Govindaraju propose the Fuzzy Run Length Smoothing Algorithm (RLSA) as a bottom-up method for the segmentation of documents. The fuzzy RLSA measure is calculated for every pixel in the input image. A new gray-scale image is created based on the RLSA measure and the image is binarized. The text line patterns are extracted from the new image. This technique has been extended to an adaptative RLSA by Gatos et al. in the sense that additional smoothing constraints are set in regard to the geometrical properties of neighboring connected components [6]. The Hough transform technique is widely used for text lines extraction. Louloudis et al. have applied this methodology to extract lines from handwritten Greek documents. In [7] and [8] the authors have proposed a methodology based on Hough transform for text lines extraction. This method has been extended in [9] to a new line/word segmentation method. Another bottom-up method based on Artificial Intelligence was

proposed in [10] by Nicolas et al. With this technique, the authors try to cluster the connected components of the document into homogeneous sets that correspond to the text lines of the document. A recent methodology makes use of adaptive local connectivity map technique for retrieving text lines from the handwritten historical manuscripts has been presented by Shi et al in [11]. In this method, the extraction of the text lines is done by a connected component collection and grouping after binarization of the input image with an adaptive binarization method. A new method proposed by Vasant Manohar et al in [12], this method involves grouping text lines segmented by a set of methods for segmentation of handwritten text lines in an undirected graph. The graph nodes correspond to connected components and the edge connecting pairs of connected components. Some important works are based on a hybrid method to extract text lines. Among these technique we can cite the method presented in [13] by Li et al. The authors use a Gaussian window and a level set method to detect text line boundaries. In [14] Bukhari et al. use the active contours (snakes) to segment handwritten text into lines. First, they apply a filter bank to smooth the input text image. The central line of text line components is then computed using ridges over the smoothed image. As a final step, active contours evolve over the ridges to segment text lines. Two unconstrained handwritten text segmentation methods are presented in [15] and [16]. The two methods propose also possibilities to apply their methods on any language. The first proposition is bottom-up method. It starts from gray values of pixels. However, the second method is top-down. It starts from overlapping blocks and use a skew blocs estimation.

### 2.2. Discussion of existent methods

Some document segmentation works estimation are based on projection [1], [2], [3], [4]. This method is very sensitive to slant, so needs some preprocessing steps such slant detection. The elimination of the slant can add distortions to the document content. Other works use some image processing tools such as binarisation, smoothing and other specific algorithms [6], [14], [15]. Third set of document segmentation methods use Hough transform such as [7], [8], [9]. The main problem of these methods are also slant. Other variety of methods based on structural description [11], [12], Artificial intelligence[9]. They try to combine image processing and pattern recognition tools [5], advanced algorithm, intelligent systems and some heuristics. The proposed method is situated in this field. It combines image processing tool and an advanced algorithm which is insensitive to slant and to size script variation.

## 3. Proposed Method

The proposed methodology for text line segmentation in handwritten Arabic document uses the algorithm of construction of the Outer Isothetic Cover of a digital object (OIC) [17]. The OIC algorithm allows



constructing the polygon of each detected connected component. So, from the information relating to each polygon and the specificity of the Arabic script, we can associate each detected component to its appropriate text line. This approach has been used for the segmentation of Bengal text in words and baseline extraction by Sark in 2010 [18]. It is not used on Arabic script processing.

### 3.1. The Arabic Script

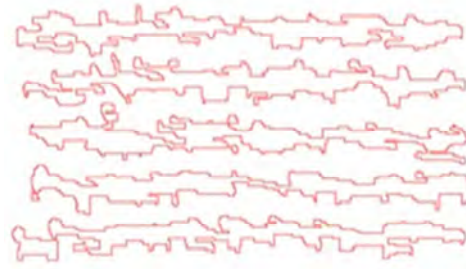
In handwritten or printed Arabic script, the Arabic word can be composed of more than one part called PAW "Pieces of Arabic Words" or also pseudo-word. A PAW is a sequence of entirely interconnected letters. A word can be composed by one or more PAWs. Another specificity of the Arabic script is that it contains complementary signs called diacritical marks. It is a secondary component of a letter which comes to complete it or to modify the sense. It is about vowels, points or other signs (chadda, madda, hamza). The number of these points can be one, two or three and having generally the same shape. As a result, Arabic text line may be considered as a set of "neighbors" PAWs and diacritical points that are cited either above or below the text line. With this hypothesis and by applying the OIC algorithm on the binary image we try to associate each detected component to its appropriate text line.

### 3.2. The Outer Isothetic Cover: OIC

The isothetic cover of a digital object specifies a simple representation of the object and provides approximate information about its structural content and geometric characteristics. When, the cover tightly encloses the object, it is said to be an outer isothetic cover (OIC). The OIC is defined by a set of isothetic polygons, having their edges lying on the grid lines, given that the effective area corresponding to the object is minimized. The algorithm of construction of the OIC of a digital object is given by Biswas et al. in [17]. By the variation of the grid size value, the authors have presented different shapes of the OIC of the treated object. The Figure 1 shows an example of the set of outer polygons for different handwritten Arabic lines. The outer isothetic covers are obtained for grid size  $g=2$ .



(a)



(b)

Figure 1. (b) Set of polygons representing handwritten Arabic lines corresponding to results of MM of (a)

Then, to draw multiple polygons corresponding to different lines, we modify the algorithm given by [17]. This idea was used by Sarkar for words segmentation of handwritten Bangla documents in [18]. So, the algorithm is applied on the result of MM on handwritten Arabic document. With a proper grid size each polygon corresponds to a single line in the initial document. In fact, on the result image of MM the grid points are traversed in the row-major order until a 90 vertex (start vertex) is found [17]. Subsequent grid points are classified, marked as "visited", and the direction is determined from each grid point to the start vertex. Finally, the OIC is constructed when the start vertex is reached again. Figure 2 presents an example of construction of OIC for the handwritten Arabic text line after application of MM. In this figure each vertex is represented by a red point, the start vertex is surrounded (green circle) and the OIC of a line is represented with blue color. With this process we extract text lines from a document image. In Figure 3 we show a simple example of text lines extraction from handwritten Arabic document.

نقطة فرقاً بين أمراض واحوال تعرض لآفات التنفس المفل

(a)



(b)



Figure 2. Construction of the OIC of handwritten Arabic lines. (a) original image, (b) result of OIC algorithm.



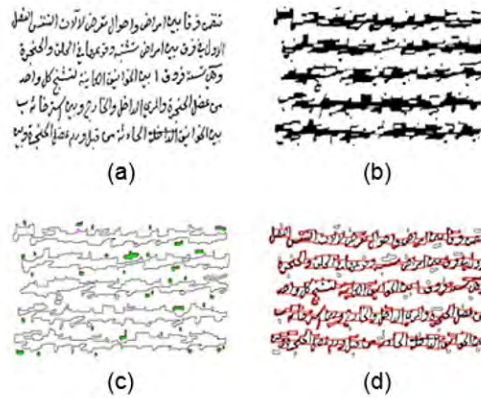


Figure 3. Text lines extraction with OIC method. (a) original image, (b) preprocessing result, (c) OIC algorithm result, (d) text lines extraction.

### 3.3. Construction of the Outer polygons of Components

The detection of each component in the document is done by the determination of its correspondent outer polygon. Applied on the binary image, the OIC algorithm allows constructing each polygon in the order of apparition of the component. Indeed, the algorithm traversing the grid points imposed on the binary image in the row major order until "start vertex" is detected (the first black pixel situated in top left) [17]. Then the construction of the first polygon is determined when the start vertex is reached again. Then the algorithm determines the polygon of the second detected component by its start vertex. The procedure is iterative and it is similar to the other components in function of their apparition in the image. Figure.1 presents an example of construction of the outer polygon of three successive components. In the Figure.4 (a), the polygon of the rectangle will be constructed in the first iteration, then the triangle polygon and in the end the polygon of the circle. In the Figure.4 (b) the polygon of the circle will be constructed in the first iteration, contrary, in the Figure.4 (c) the polygon of the triangle will be the first detected.

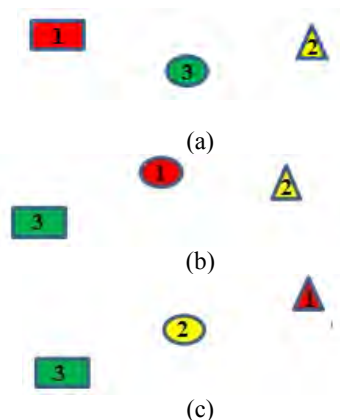


Figure 4. Order of Construction of the Polygons

### 3.4. Assign Detected Component to Text Line

As we mentioned above, Arabic text is composed of a set of

PAWs and diacritic points. In function of the perimeter of the given polygon, we can identify a PAW from a diacritic point. Generally, a diacritic point is presented by a small polygon.

On the other hand, with the detected polygon we can determine the extremities of the connected component. Indeed, the polygon is represented by its vertices of coordinates of pixel  $(i, j)$  in the image. Let  $\{(x_i, y_i)\}$  ( $i=1..n$ ,  $n$ : number of vertices) the set of the coordinates of vertices of the polygon. The maximum (respectively the minimum) of the  $\{x_i\}$  represents the upper limit (respectively the lower limit) of a connected component. Thus, perimeter, upper and lower limit of each detected component are important information that will help us to associate the component to its appropriate text line.

#### 3.3.1. Association of PAW to its Appropriate Text Line:

By comparing the value of perimeter of the polygon of the detected component to a value  $parm$  allows us to identify a PAW compared to a diacritic point. Two existing case for associate a PAW to text line. The process is as following.

- Case 1: PAW can be detected in the first iteration of the application of the OIC algorithm. In this particular case, the PAW is affected to the first text line in the treated document. The PAW will be labeled by a color  $C_k$ . Figure.5 shows an example of this case.



Figure 5. Case when The First Detected Component is a PAW

- Case 2: PAW is detected after a labeled PAW or a labeled diacritic point. In the Figure.6, the current PAW is detected after a labeled diacritic point (Figure.6 (a)), so, it will be labeled with the same color since it verifies the property (1c) (Figure.6(b)). In this case, the current PAW will be decided belonging to the line of the previously detected component or not, in function of its limits and those of its predecessor. We have proposed three properties that can verify this situation. Indeed, let  $minPAW_c$  (respectively  $maxPAW_c$ ) the upper limit (respectively the lower limit) of the current PAW and let  $minPAW_p$  (respectively  $maxPAW_p$ ) the upper limit (respectively the lower limit) of its predecessor detected PAW. Whether the property (1a) or (1b) or (1c) is verified, the current PAW is assigned belonging to the same text line of the predecessor PAW and it will be labeled with the color  $C_k$  of its predecessor. Subsequently the same process is done for the other successive detected PAW until the properties (1a), (1b) and (1c) are not verified. In this step, we attribute another color  $C_{k+1}$  to the PAW that does not verify those properties and we start constructing the new text line with the same principle.





Figure.7 illustrates an example for the case where the current PAW is preceded by another PAW. As shown in this figure, the first detected paw is the second in the handwritten word (Figure.7(a)) and it will be the first labeled (Figure.7(b)). Then, the third paw is detected in the second iteration of the algorithm (Figure.7(b)) and it will be labeled with the same color of its predecessor (Figure.7(c)) since it verifies the two properties (1a) and (1c).

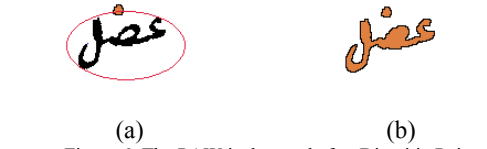


Figure 6. The PAW is detected after Diacritic Point

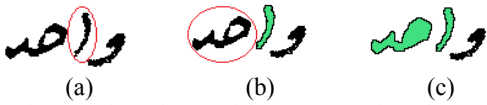


Figure 7. The Predecessor of the Current PAW is another PAW

In the Figure.8, we illustrate an example of text line extraction. Figure.5(a) presents the initial image. Figure.8(b) shows the detection of the current PAW (circled in red) after the extraction of the first text line. For this current PAW neither property (1a), nor (1b), nor (1c) is verified. In this case, we attribute a new color for this PAW (Figure.8(c)) and we start to construct a new text line. The final result is shown in the Figure.8(d).



Figure 8. The PAW is detected after PAW

$$\min PAW_p \leq \min PAW_c \leq \max PAW_p \quad (1a)$$

$$|\max PAW_c - \max PAW_p| \leq S$$

$$\min PAW_p \leq \max PAW_c \leq \max PAW_p \quad (1b)$$

$$|\min PAW_c - \min PAW_p| \leq S \quad (1c)$$

Where the threshold value  $S$  is selected and fixed so that the current PAW and its previous are nearest in the sense of their limits.

**3.3.2. Association of Diacritic Point to its Appropriate Text Line:** Diacritic point is a connected component that is cited above or below the PAW. On the other hand, small polygons extracted with the OIC algorithm generally correspond to the diacritic point or punctuation mark. We identify these components by comparing their perimeter to a chosen value  $parm$ . In our case, the value  $parm$  is fixed to 45 after a learning phase. Based on these two hypotheses, we try to associate

diacritic point to its appropriate text line.

Diacritic point cited above PAW: two different cases can exist:

**Case 1:** On the first iteration, the connected component detected by the OIC algorithm is a diacritic point (Figure.9 (a)). So, the point will be affected to the first text line and will be labeled with the first color  $C_k (k=1)$  (Figure.9 (b)).

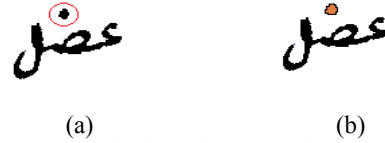


Figure.9. The First detected component is a Diacritic Point

**Case 2:** The diacritic point is detected after one or more PAWs. In this case, the point will be associated to the same text line of its previous components if it verifies the property (2), where  $\min PD$  represents the upper limit of the diacritic point and  $\{\min PAW_i\}$  represents the set of the upper limits of the  $i$  previous PAWs. Figure.10 shows an example of this case.

$$\min PD \geq \min \{\min PAW_i\} \quad (2)$$



Figure. 10. The upper limit of the diacritic is equal to or situated after the upper limit of the precedent paw.

•Diacritic point cited below PAW: Two properties to verify for deciding that diacritic point either belonging to same text line of its previous components or else it belonged to the next text line.

**Case 1:**

$$\min PD \leq \max \{\max PAW_i\} \quad (3)$$

Where  $\min PD$  represents the upper limit of the diacritic point and  $\{\max PAW_i\}$  represents the set of the lower limits of the  $i$  previous PAWs.

**Case 2:** if the precedent property (3) does not satisfy, we verify, if the upper limit of the diacritic point is close to the maximum of the set of lower limits of its previous components (i.e. it verifies the property (4)).

$$|\min PD - \max \{\max PAW_i\}| \leq S \quad (4)$$

Where the threshold value  $S$  is selected and fixed after a number of learning test.

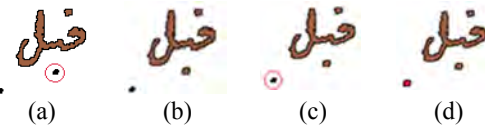


Figure. 11. Two Different Cases of Labeling of Diacritic Point

For the diacritic point cited below PAW (Figure.11(a,c)), it will be decided belonging to the same text line of the PAW if it verifies respectively (3) or (4), in this case it will be labeled with the same color  $C_k$  of the previous PAW



(Figure.11(b)). If neither (3) nor (4) is satisfied, the point will be labeled with the color  $C_{k+1}$  of the next text line (Figure.11(d)).

#### 4. Experimental results

The proposed text line extraction method is tested on a set of 100 randomly chosen handwritten Arabic text images selected from KHATT Database [20]. The khatt Database is a collection of 1000 handwritten forms written by 1000 distinct writers from different countries. The database contains also 2000 randomly selected text paragraphs from 46 sources, 2000 minimal text paragraph covering all the shapes of Arabic characters, and optionally written paragraphs on open subjects. The 2000 random text paragraphs consist of 9327 lines. The database forms were randomly divided into 70%, 15%, and 15% sets for training, testing, and verification, respectively. The sub-set of the 100 handwritten Arabic text images are chosen from the set "Test" of KHATT Database. This sub-set text images are written by different writers. Some of them have variable skew angles among text lines. In addition, there are text images having text lines with different skew directions as well as text images having text lines with converse skew angles along the same text line. The total number of text lines is 590. Manual evaluation of the results of our method on these text images shows a correct extraction rate of 74%. Figure.12 illustrate an example of correct text line extraction for a text written by different writers.

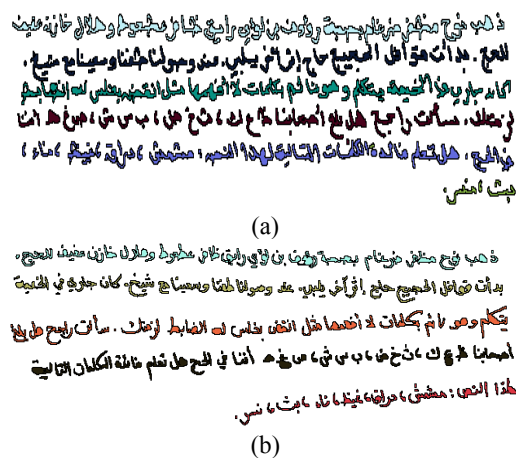


Figure. 12. Correct Text Line Extraction. (a) Text with small interline and (b) Text with skewed and curved text lines

The weak correct extraction rate is explained by different factors. Indeed, it might be due to the choice of the two threshold  $S$  and  $parm$  for associating connected components to their line. the second factor is explained by the fact that a component of a line ( $i + 1$ ) may be detected before some components of the line ( $i$ ), this case is almost presents in case of two adjacent curved or skewed lines. Another factor is due to the touching and overlapping of certain lines. Finally, it can be explained by the small number of the used text images for the evaluation. Figure.13 shows an example of incorrect text line extraction. This work is actually evaluated on the

HAIJ database [27] created to improve our actual research on the innovated integrated intelligent and mobile system for tutorial and guiding of HAIJ steps based on geolocalization.

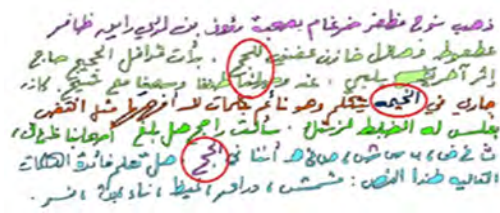


Figure. 13. Incorrect Text Line Extraction

#### 5. Comparison with existent methods

The comparison with existent methods is done essentially on the size of data evaluation. The obtained results can be compared to presented method only if the evaluation is done on the same datasets and by the use of the same evaluation method. In table II, we present some obtained results. Most of these methods do not indicate the origin of their document. Even if they indicate the name of the used database, they do not indicate how they choice the training set and the testing set. They only give an idea about the size of their dataset. The proposed method gives the worst segmentation rate. Firstly, we should do comparison on the same dataset. Secondly, improvement of the segmentation method and or the evaluation method is also needed.

TABLE II. SEGMENTATION METHODS COMPARISON

Approaches	Database Language	Pages	Line number	Segmentation rate
Arivazhagan [1]	English and Arabic	720	11581	97.31
Zahour [2]	Printed or handwritten historical Arabic	100		96
Nicolaou [4]	Latin	80	1771	98.8
Louloudis [10]	Latin	152	3382	95.8
Shi and al [14]	DARPA MADCAT database *	45	1022	99.5
Alaei et. Al. [18]	Persian text	52	823	92.35
OIC approach	Unconstrained Arabic handwriting KHATT database	100	590	74

\* blank paper, paper with pre-printed ruled lines and letterheads with company logos

#### 6. Conclusion

We propose a new bottom-up method based on the algorithm of construction of the outer isothetic cover (OIC) for handwritten Arabic text line segmentation. By comparing respectively the upper and lower limits (respectively the perimeter) of two consecutive detected components to a threshold, the method allows to associate PAW or diacritic point to its appropriate line. This method can extract skewed, slightly overlapping and curved text lines.

In many cases, the algorithm of the construction of the OIC can detect the components of the line  $i + 1$  before these of line  $i$ , in this case, the components of  $i$  will be associated to the line  $i + 1$ . In the scope of feature work, we consider





developing a processing step that allows solving and overcoming this problem. We will focus also to test our method on a larger number of image documents such as the HAJJ database and we will aim to evaluate the performance of the proposed method by counting the number of matches between the detected text lines by our method and the text lines in the ground truth by using the MatchScore Table. The main perspective of this work is to be able to extract correctly information from Arabic scanned documents. The extracted information will be transformed to text with OCR. Later, the obtained text will be analysed by an intelligent system to extract HAJJ rules that will be integrated to a mobile system for tutorial and guiding of HAJJ steps based on geolocalization (GPS).

## 7. Acknowledgements

This research was supported by the strategic Technologies Programs of the National Plan or Science, Technology and Innovation (MAARIFAH) in the Kingdom of Saudi Arabia. No: 13-INF134-05.

## 8. References

- [1] M. Arivazhagan, H. Srinivasan, and S. N. Srihari, "A statistical approach to handwritten line segmentation," in Document Recognition and Retrieval 14, Proc. of SPIE, San Jose CA, USA, pp. 65 000T.1–65 000T.11, 2007.
- [2] A. Zahour, L. Likforman-Sulem, W. Boussalaa, and B. Taconet, "Text line segmentation of historical arabic documents," in Proc. of the International Conference on Document Analysis and Recognition (ICDAR), Curitiba, Paran, Brazil, pp. 138–142, 2007.
- [3] F. Shafait, D. Keysers, and T. M. Breul, "Performance comparison of six algorithms for page segmentation," IAPR Workshop on Document Analysis Systems, vol. 3872, no. 1, pp. 368–379, 2006.
- [4] A. Nicolaou and B. Gatos, "Handwritten text line segmentation by shredding text into its lines," in Proc. of the International Conference on Document Analysis and Recognition, Barcelona, Spain, pp. 626–630, 2009.
- [5] Z. Shi and V. Govindaraju, "Line separation for complex document images using fuzzy runlength," in Proc. of the International Workshop on Document Image Analysis for Libraries (DIAL), Washington, DC, USA, p. 306, 2004.
- [6] B. Gatos, A. Antonacopoulos, and N. Stamatopoulos, "Line separation for complex document images using fuzzy runlength" in Proc. of the International Conference on Document Analysis and Recognition (ICDAR), Curitiba, Brazil, pp. 1284–1288, 2007.
- [7] Fadoua Bouafif, Samia Snoussi Maddouri, Nouredine Ellouze, On Segmentation Methods for multilingual and mixed Scripts, ICGST-GVIP journal emphasizes on graphics, vision and image processing, 2009.
- [8] G. Louloudis, B. Gatos, and I. Pratikakis, "Text line detection in hand-written documents," PATTERN RECOGNITION, vol. 41, pp. 3758–3772, 2008.
- [9] G. Louloudis et al., "Text line and word segmentation of handwritten documents," PATTERN RECOGNITION, vol. 42, pp. 3169–3183, 2009.
- [10] S. Nicolas, T. Paquet, and L. Heutte, "Text line segmentation in handwritten document using a production system," in Proc. of the International Workshop on Frontiers in Handwriting Recognition (IWFHR), Tokyo, Japan, pp. 245–250, 2004.
- [11] Z. Shi, S. Setlur, and V. Govindaraju, "A steerable directional local profile technique for extraction of handwritten arabic text lines," in Proc. of the International Conference on Document Analysis and Recognition (ICDAR), Barcelona, Spain, pp. 176–180, 2009.
- [12] V. Manohar, S. N. Vitaladevuni, C. H. Cao, R. Prasad, and P. Natarajan, "Graph clustering-based ensemble method for handwritten text line segmentation," in Proc. of the International Conference on Document Analysis and Recognition (ICDAR), Beijing, China, pp. 574–578, 2011.
- [13] Y. Li, Y. Zheng, D. Doermann, and S. Jaeger, "Script-independent text line segmentation in freestyle handwritten documents," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, pp. 1313–1329, 2008.
- [14] S. S. Bukhari, S. F., and T. M. Breuel, "Script-independent handwritten textlines segmentation using active contours," in Proc. of the International Conference on Document Analysis and Recognition (ICDAR), Barcelona, Spain, pp. 446–450, 2009.
- [15] A. Alaei, P. Nagabhushan, Umapada Pal: Piecewise painting technique for line segmentation of unconstrained handwritten text: a specific study with Persian text documents. Pattern Anal. Appl. 14(4): 381–394, 2011.
- [16] M. Ziaratban, K. Faez: Adaptive Script-Independent Text Line Extraction. IEICE Transactions 94-D(4): 866–877, 2011.
- [17] A. Biswas, P. Bhowmick, and B. Bhattacharya, "Construction of isothetic covers of a digital object: A combinatorial approach," Journal of Visual Communication and Image Representation, vol. 21, pp. 295–310, 2010.
- [18] A. Sakar, A. Biswas, P. Bhowmick, and B. Bhattacharya, (2010). Word segmentation and baseline detection in handwritten documents using isothetic covers. In ICFHR10, pages 445–450, 2010.
- [19] S. A. Mahmoud, I. Ahmad, M. Alshayeb, W. G. Al-Khatib, M. T. Parvez, G. A. Fink, V. Margner, and H. E. Abed, "Khatt: Arabic offline handwritten text database," in Proc. of the International Conference on Frontiers in Handwriting Recognition (ICFHR), Bari, Italy, Sep, pp. 447–452, 2012.
- [20] I. Phillips and A. Chhabra, "Empirical performance evaluation of graphics recognition systems," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 9, pp. 849–870, 1999.
- [21] Y. Alginahi, A survey on Arabic character segmentation, International Journal on Document Analysis and Recognition (IJDAR), Volume 16, Pages: 105–126, 2013.
- [22] A. ElNagar, R. Bentrcia, A Multi-Agent Approach to Arabic Handwritten Text Segmentation, International Journal of Advanced Research in Computer Science and Software Engineering, , Volume 4, Pages: 207–215, 2012.
- [23] A. T. Jamal, N. Nobile, and C. Y. Suen, Shape Recognition for Arabic Handwritten Text Segmentation, Artificial Neural Networks in Pattern Recognition, Canada, Volume 8774, Pages: 228–239 2014.
- [24] N. Aouadi, S. Amiri and A. Kacem Echi, Segmentation of Connected Components in Arabic Handwritten Documents, International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA), Pages: 738–746, 2013.
- [25] A. M. Zeki, M. S. Zakaria and C-Y Liong, Segmentation of Arabic Characters: A Comprehensive Survey, Technology Diffusion and Adoption: Global Complexity, Global



- Innovation, chapter 16, 2013.
- [26] A. T. Jamal, N.Nobile, and C.Y.Suen, Shape Recognition for Arabic Handwritten Text Segmentation, Artificial Neural Networks in Pattern Recognition, Canada, Volume 8774, Pages: 228-239 2014.
  - [27] Y.Wahabi, S. Snoussi Maddouri, HAJJ and Umrah digital library for the preservation and documentation of references and Arabic manuscripts, Arabic language and modern softwares conference, AlJouf University, Saudi Arabia, 2014



## Biographies



**Samia Snoussi Maddouri**, received her diploma (Dipl.-Ing.) from the faculty of science of Tunis in 1995, Master in system and signal processing and doctorate (Dr.-Ing.) degrees in electrical engineering from the National engineering school of Tunis (ENIT), in 1997 and 2003 respectively. Since 1995, she has been working as university teacher in different Tunisien institutes. Currently she is a member of the research staff in the Image and signal processing laboratory at ENIT and teaching staff at faculty of computer science and information technologies at Jeddah university in Saudi Arabia. She worked as chef of computer science department in Taiba university. Her main areas of research are image processing and pattern recognition. Currently, she is working on Document segmentation and analysis methods and Arabic script recognition. She developed recognizer for handwritten words of Tunisian town names of IFN/ENIT database. Structural global primitives and local Fourier descriptors with a Transparent Neural Network based recognizer are the key features of her solution. Since 1998 she is also working on Document segmentation methods, at the beginning on bank checks and then on different text documents printed and handwritten. The distinction between Arabic and Latin handwritten and printed script is also one on her area of research. These Works are done in close cooperation with Braunschweig Technical University, Germany. She is one of the first developers of the IFN/ENIT-Database in 2002 (used by more than 100 research groups from more than 30 countries). She published more than 40 papers including journal papers. She was a member of DAAD cooperation project "on the Way to the Information Society: Branshweig-Tunis" during 3 year and participates to the organization of workshops in this field and represents the members of her laboratory and the Tunisian young academics women in the DAAD meeting of Project Representatives in the Programme German-Arab/Iranien Higher Education Dialogue: Dialogue through Cooperation at the Technical University, Germany in 2008. Actually I am a member of a the project intitlled "Innovated Integrated Intelligent and Mobile System for tutorial and guiding of HAJJ steps based on geo-localization (GPS)".



**Fethi Ghazouani**, received his diploma (Master's degree in Computer Science) from the Faculty of Sciences of Tunis (FST), Tunisia in 2005 and he has received his Master degree in Computer Sciences (Data, Knowledge and Distributed Systems) from the University of Jendouba (FSJEG), Tunisia, in 2013. He has part of university study of the skills in programming, in data mining, in image processing, in

artificial learning, oriented databases, etc. In his master study, he has working on Document segmentation and analysis methods and Arabic script recognition. He has developed a system for segmentation of Arabic documents that allows segmenting the document into text lines, words and characters. This work is done in cooperation with the IfN institute in Braunschweig University, Germany in 2011. Currently, he is working toward the Ph.D. degree with the Ecole Nationale des Sciences de l'Informatique (ENSI), University of Manouba, Tunisia. He is also a Permanent Researcher at RIADI Laboratory, University of Manouba, since 2014. His work is mainly related with machine learning and knowledge modeling applied to the remote sensing images.



**Yosra Wahabi** has the License in Applied Computer Sciences to the Management from Economic Sciences and Management faculty of Nabeul in Tunisia after presenting a research about a Database Management System of arabic document images collaborated with the Image and signal processing Laboratory of the National Engineering School in Tunis (ENIT). She has the research Master degree in signal processing with the same laboratory in 2007. Her main areas of research are image processing, object tracking and pattern recognition. Actually she is working on document segmentation, analysis methods and Arabic script recognition. In 2008 she has the certificate of skill in computer science C2i, Master Teacher from the Educational ministry in Intel program. She has research training in March 2009 with the INRIA group, Sophia Antipolis in France. In the same year she provides training workshops in computer science for teachers of other disciplines working in the educational ministry of Tunisia. She worked as computer science teacher in the same ministry for six years correlated with many trainings in the pedagogy field. Since 2012 she is a teacher in Umm Al Qura University, Saudi Arabia. She published three papers: Arabic language and modern softwares conference, Aljouf, Saudi Arabia in 2014, The International Arab Journal of Information Technology (IAJIT) in 2010 and the International Conference on Image Processing, Computer Vision, & Pattern Recognition (ICCV), USA in 2008. Since 2015 she is a member of the "Innovated Integrated Intelligent and Mobile System for tutorial and guiding of HAJJ steps based on geo-localization (GPS)" research project.



